CLASSIFICATION WITH DEPENDENT TRAINING SAMPLE*

by

S. Bandyopadhyay

Technical Report No. 230

August, 1974

University of Minnesota
Minneapolis, Minnesota

* Partly supported by U.S. Army Research Grant DA-ARO-D-31-124-70-G-102 at the University of Minnesota, Minneapolis, Minnesota.

ACKNOWLEDGEMENT

The author wishes to express his gratitude to Professor Somesh Das Gupta for providing guidance and continuing encouragement during the course of this research.

TABLE OF CONTENTS

ABSTRACT

An observation vector $X:p\times 1$ is available on an experimental unit which is a random outcome from a population π. It is known that π is identical to one of two specified populations π_1 and π_2, which represent the same population at two different states or points of time. The information on the distribution of X is obtained from a training sample of size N, each observed at these two points of time; let these values be denoted by $X_{1\alpha}$, $X_{2\alpha}$, respectively, $(\alpha=1,\ldots,N)$. It is assumed that $X \sim N_p(\mu_i, \Sigma)$ when the unit comes from π_i. Moreover, it is assumed that $(X_{1\alpha}, X_{2\alpha})$ is distributed as 2p-variate normal distribution. Different cases are considered according as the parameters are known, partially known, or unknown. Estimated and plug-in likelihood ratio rules are defined and derived and the exact distributions of the statistics involved are obtained. Admissibility of some of these rules are shown. Following Anderson (Discriminant Analysis and Applications. Ed. T. Cacoullos. Academic Press, 1973), asymptotic expansions of the distributions of these statistics, properly normalized and studentized, are derived for large N. The probabilities of correct classification (PCC) have been studied with respect to the dependence between $X_{1\alpha}$ and $X_{2\alpha}$, and moreover these probabilities are shown to be increasing in Mahalanobis distance between the two populations when $\mathrm{Cov}(X_{1\alpha}, X'_{2\alpha}) = \rho\,\Sigma$. Some results regarding estimates of these PCC's and inequalities involving their expectations are obtained.

1. INTRODUCTION AND NOTATIONS.

1.0. Introduction.

The problem of classification deals with one or several units each to be classified in one of k different well-identified populations. Since it is not known, in which one of the k populations a unit belongs, a vector X of p measurements or characteristics on that unit is used for classification. The probability distribution of this vector of p variables are assumed to be different in these k populations. Wald [28], Anderson [2] and Rao [21] considered the classification problem when the distributions of X in k populations are known and, in particular, when these k distributions are multivariate normal with the same covariance matrix. Often it is assumed that the populations belong to a parametric family with some or all the parameters being unknown; then the unknown parameters are estimated using correctly classified samples from those k populations. These samples are sometimes called "training samples". Wald [28], Anderson [2] and Rao [21] suggested some heuristic rules for classification when the k distributions are multivariate normal with unknown means but the same (known or unknown) covariance matrix. For a lengthy review of work in this area, see DasGupta [10] and the bibliography by Cacoullos and Styan [8].

It happens sometimes that the k populations represent the same population at k different points of time or at k states. In order to get information on the unknown distribution functions we shall take a random sample from this population and make p measurements on each of the sampled units at each of these k specified time-points. Thus the k samples originating from these k populations become dependent since they are based on the same set of units. We shall call such a training sample a "dependent training sample". In the usual design, the k samples are based on independently drawn units from k populations. It appears that there is no literature on the classification problem based on a dependent training sample.

We shall illustrate our problem by giving an example. Suppose an employer has two sets of scores of a certain test taken by a group of persons before and after a training program. A prospective employee is given the same test and the employer wants to use the new score together with the two sets of scores he has to decide whether or not the candidate needs the training. Now, the employer faces a problem since the standard classification techniques are not applicable. This is because the distribution of the two sets of scores may be pairwise dependent as they are based on the same group of individuals. The above situation is a problem of classification with dependent training samples.

We shall consider this problem for two populations and introduce mathematical notations to make it precise. Our results may be generalized for $k > 2$ proceeding as in the standard case.

Let ω be an experimental unit which is a random outcome from a population π. It is known that π is identical with one of the two specified populations π_1 and π_2. We shall consider a situation when Π_1 and Π_2 denote the same population at two different states or points of time t_1 and t_2. Let $X = X(\omega)$ be a $p \times 1$ vector of observation on ω and the distribution function of X be F_i when $\Pi = \Pi_i$, $i = 1,2$. The problem is to identify Π with one of Π_1 and Π_2 on the basis of X and the knowledge of F_1 and F_2. When F_1 and F_2 are not completely known, we shall get information about them based on a random sample of N units $\omega_1, \omega_2, \ldots, \omega_N$ with $X_{i\alpha}$ as the X observation on the unit ω_α, observed at time t_i. $\{X_{i\alpha};\ i = 1,2;\ \alpha = 1,2,\ldots,N\}$ constitutes our dependent training sample.

Consider a situation when the distribution of X from Π_i is $N_p[\mu_i, \Sigma_{ii}]$, $i = 1,2$, and $(X'_{1\alpha}, X'_{2\alpha})'$, $\alpha = 1,2,\ldots,N$ are independently distributed as

$$N_{2p}\left[\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix},\ \Gamma\right]$$

where

$$\Gamma = \begin{pmatrix}\Sigma_{11} & \Sigma_{12}\\ \Sigma_{21} & \Sigma_{22}\end{pmatrix}$$

and where the μ_i's are $p \times 1$ vectors and Σ_{ij}'s are $p \times p$ matrices such that Γ is positive definite. For most of our work we shall assume block-circularity in Γ, i.e., the covariance matrix of $X_{i\alpha}$ and $X_{j\alpha}$ is given by

$$\text{Cov}\,[X_{i\alpha}, X'_{j\alpha}] = \Sigma_{|i-j|} \quad (i,j = 1,2)\ ;$$

so, we have $\Sigma_{11} = \Sigma_{22} = \Sigma_0$, $\Sigma_{12} = \Sigma_{21} = \Sigma_1$. The 'standard' two population classification problem is a special case of the above set-up, namely, when Σ_1 is a zero matrix. The following model motivated the above structure of Γ. Let X_t be the X observation, observed at time t and such that

$$X_t = m_t + U_t$$

where $\{U_t\}$ is a stationary Gaussian process with discrete time parameter (Anderson [2], p. 373) i.e.,

$$U_t \sim N_p[m, \Sigma]$$

and

$$\text{Cov}\,[U_t, U_s'] = \Sigma_{|t-s|} .$$

Examples of physical situations, where Γ is a block-circular matrix, may be found in Olkin and Press [20] and references there in. A special case of block-circular Γ will also be treated by considering

$$\Sigma_1 = \rho\Sigma_0 , \quad |\rho| < 1 .$$

The above structure of the covariance matrix Γ arises from a process which is a special case of a stationary Gaussian process. Consider a first order autoregressive process

$$U_t = \lambda\, U_{t-1} + \epsilon_t, \qquad t = 0, \pm 1, \pm 2, \ldots$$

where ϵ_t's are independently distributed as $N_p[0, \Lambda]$. It is known that (Anderson [2], p. 166) for every t and r and for $|\lambda| < 1$,

$$\begin{pmatrix} X_{t-r} \\ X_t \end{pmatrix} \sim N_{2p}\left[\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \Sigma & \rho\Sigma \\ \rho\Sigma & \Sigma \end{pmatrix}\right]$$

where

$$\mu_1 = m_{t-r}, \qquad \mu_2 = m_t$$

and

$$\Sigma = (1-\lambda^2)^{-1}\Lambda, \quad \rho = \lambda^r .$$

π_1 and π_2 may be considered as any two time points $t-r$ and t.

We shall always consider the mean vectors to be unknown and deal with several cases arising out of specified knowledge and structure of Γ.

In Section 2, the estimated likelihood ratio (ELR) rules and the plug-in likelihood ratio (PLR) rules are defined and derived. The concept of ELR rules are due to Anderson [1]. The ELR rules are easily obtained when Γ is known or no structure of Γ is assumed when Γ is unknown. When Γ is unknown and block-circularity in Γ is assumed, the ELR rule does not have a closed form. However, it is possible to give a closed form of the ELR rule when partial knowledge about block-circular Γ is assumed. The notion of PLR rules was first introduced by Wald [27]. The PLR rules are much simpler to compute and we have obtained closed forms in all cases (except that, in one case we have assumed $p = 1$). In sections following

Section 2, we have considered block-circular Γ only.

In Section 3, we derive the density functions of the statistics involved in the classification rules, considered in Section 2, which have closed forms. The distribution aspect in the standard case has been considered by John [14] when Σ_0 is known and by Sitgreaves [24] when Σ_0 is unknown. Two different representations of the classification statistic are available for the standard case, one is given by Wald [27] and the other is by Bowker [6]. We suitably modify our problem to use the methods of John [14] and Sitgreaves [24].

In Section 4, the asymptotic expansions of the distribution functions of the statistics considered in Section 3 are obtained for large N to the order of N^{-2}. In the standard case John [15] has obtained the asymptotic expansion of the distribution function involved in the PLR rule when Σ_0 is known, to the order of N^{-2}; Σ_0 unknown case was considered by Okamoto [19] and Bowker and Sitgreaves [7], to the order of N^{-3}. In a recent paper, Anderson [5] has also considered the asymptotic expansion, in the standard case, using a rather simple method. John [15], Okamoto [19] and Bowker and Sitgreaves [7] have normalized the classification statistic using the asymptotic mean and the asymptotic variance of the statistic whereas Anderson [5] normalized the classification statistic using consistent sample estimates of the asymptotic mean and the asymptotic variance of the statistic. We have followed Anderson's [5] method of expansion to find the asymptotic expansions for the normalized classification statistics (normalized using the asymptotic mean and variance and also their consistent sample estimates) and, as a special case, obtained approximation on the probabilities of correct classification (PCC) to the order of N^{-2}.

In Section 5, we studied the PCC of some classification rules, their estimates and inequalities involving them. For the standard case, the PCC of a class of classification rules has been shown to be monotonically increasing with $\alpha \equiv (\mu_1 - \mu_2)' \Sigma_0^{-1}(\mu_1 - \mu_2)$, by Das Gupta [11]. We shall show a similar result. We shall also study the PCC in detail, numerically, when $p = 1$, as a function of ρ. In the standard case with Σ_0 known and $p = 1$, Hills [13] has established inequalities concerning the PCC's of the ELR rule, LR rule when parameters are known and expected values of Fisher's [12] estimate and Smith's [25] estimate. Das Gupta [11] has generalized Hill's result to the general case of $p \geq 1$ and Σ_0 being known or unknown. Sorum [26] and Lachenbruch [18] have studied the problem from another point of view. We have established the same inequalities as obtained by Das Gupta [11] in two different cases, viz. Σ_0 is known and Σ_0 is unknown but $\Sigma_{12} = \rho\,\Sigma_0$ with known ρ. Exact distributions are obtained for $p = 1$.

In Section 6, the admissibility aspect is considered. In the standard case, Das Gupta [9] has proved the admissibility of the ELR rule when Σ_0 is known and the admissibility of the ELR rule among the class of all invariant rules when Σ_0 is unknown. Kiefer and Schwartz [16] have proved admissibility of the ELR rules in a more general framework. In our proof we use Lemma 3 of Kiefer and

Schwartz [16] and establish the admissibility of the ELR rules.

Since our model includes two population classification problems with equal sample sizes from each population, we have been able to check our findings and compare the differences with the results for the standard case.

1.1. Notations: The following notations will be used in the sequel; the other notations will be explained at the appropriate places:

(i) $\varphi(\cdot)$ and $\Phi(\cdot)$ will represent, respectively, the density function and the distribution function of a standard normal variable.

(ii) In general we shall denote a density function by $f(\cdot)$ (except for standard normal density) and by $f_i(\cdot)$ when the hypothesis H_i $(i = 1,2)$ obtains.

(iii) For a $p \times 1$ vector X and a $p \times p$ non-singular matrix B,

$$((X)) \equiv X'X$$

and

$$((X;\ B)) \equiv X'B^{-1}X$$

(iv) $W_p[\Sigma, m]$ will stand for a p-variate Wishart distribution with m degrees of freedom and parameter Σ.

(v) For a scalar a and a $p \times p$ matrix A,

$$\exp(a) \equiv e^a$$

$$\mathrm{tr}(A) \equiv \text{trace of the matrix } A$$

$$\mathrm{etr}(A) \equiv \exp[\mathrm{tr}(A)] .$$

2. CLASSIFICATION RULES.

2.0. Introduction.

Consider a $p \times 1$ vector X_0 from $N_p[\mu, \Sigma]$. It is known that $(\mu, \Sigma) \equiv (\mu_i, \Sigma_{ii})$ for exactly one i, $i = 1,2$. Let H_i denote the hypothesis that $(\mu, \Sigma) = (\mu_i, \Sigma_{ii})$, $i = 1,2$. When μ_i and Σ_{ii}, $i = 1,2$, are known, a likelihood ratio rule accepts H_1 if, and only if,

$$(2.1) \quad \left[|\Sigma_{22}|/|\Sigma_{11}|\right]^{\frac{1}{2}} \operatorname{etr}\left[-\tfrac{1}{2}\{((X_0-\mu_1; \Sigma_{11}))-((X_0-\mu_2; \Sigma_{22}))\}\right] > c$$

for some constant c.

We shall call the rule (2.1) a maximum likelihood rule when $c = 1$.

When parameters are partially or completely unknown, we shall consider a training sample, independent of X_0, of size N,

$$\{\begin{pmatrix} X_{1\alpha} \\ X_{2\alpha} \end{pmatrix}, \alpha = 1,2,\ldots,N\} \text{ from } N_{2p}\left[\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \Gamma \equiv \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}\right]$$

where $X_{i\alpha}$, $i = 1,2$ and $\alpha = 1,2,\ldots,N$ are $p \times 1$ vectors and parameters in the distributions are given by the corresponding partition.

At this point we have two ways to construct a classification rule, one based on the training sample as well as the observation to be classified and the other one based only on the training sample.

Let $L_i(\mu_1, \mu_2, \Gamma)$ denote the likelihood based on the training sample and X_0, under H_i, $i = 1,2$.

<u>Definition 2.1.</u> An estimated likelihood ratio (ELR) rule is defined as the rule which accepts H_1 if, and only if,

$$\sup L_1(\mu_1, \mu_2, \Gamma)/\sup L_2(\mu_1, \mu_2, \Gamma) > k \tag{2.2}$$

for some constant k, where sup is the supremum and is taken over the unknown parameters.

When $k = 1$, we shall call the ELR rule as the estimated maximum likelihood (EML) rule.

<u>Definition 2.2.</u> A plug-in likelihood ratio (PLR) rule is defined as the rule given by the equation (2.1) with unknown parameters replaced by their respective maximum likelihood estimates based only on the training sample.

We shall define the plug-in maximum likelihood (PML) rule as the PLR rule with $c = 1$.

We shall always consider μ_1 and μ_2 to be unknown and the following cases for Γ:

(a) Γ is known.

(b) Γ is unknown.

(c) Σ_{11} and Σ_{22} are known, Σ_{12} is unknown.

(d) $\Sigma_{11} = \Sigma_{22} = \Sigma$, $\Sigma_{12} = \Sigma_{21}$, and both Σ and Σ_{12} are unknown.

(e) $\Sigma_{11} = \Sigma_{22} = \Sigma$, $\Sigma_{12} = \rho\Sigma$, $|\rho| < 1$, and Σ is known but ρ is unknown.

(f) $\Sigma_{11} = \Sigma_{22} = \Sigma$, $\Sigma_{12} = \rho\Sigma$, and Σ is unknown but ρ is known.

(g) $\Sigma_{11} = \Sigma_{22} = \Sigma$, $\Sigma_{12} = \rho\Sigma$, and both Σ and ρ are unknown.

2.1. ELR rules.

The likelihood based on the training sample and X_0, under H_1, can be written as

$$L_1(\mu_1, \mu_2, \Gamma) = c_0 |\Sigma_{22\cdot1}|^{-N/2} |\Sigma_{11}|^{-(N+1)/2}$$

$$\cdot \exp\left[- (1/2) \sum_{\alpha=1}^{N} ((X_{2\alpha} - \mu_2 - \beta_{21}(X_{1\alpha} - \mu_1); \Sigma_{22\cdot1}))\right]$$

$$\cdot \exp\left[- (1/2) \sum_{\alpha=1}^{N} ((X_{1\alpha} - \mu_1; \Sigma_{11})) - (1/2)((X_0 - \mu_1; \Sigma_{11}))\right]$$

where

$$\beta_{21} = \Sigma_{21} \Sigma_{11}^{-1}$$

$$\Sigma_{22\cdot1} = \Sigma_{22} - \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12}$$

and

$$c_0^{-1} = (2\pi)^{p(2N+1)/2} .$$

Let

$$\bar{X}_i = (1/N) \sum_{\alpha=1}^{N} X_{i\alpha} \qquad i = 1,2$$

$$(2.3) \qquad S_{ij} = \sum_{\alpha=1}^{N} (X_{i\alpha} - \bar{X}_i)(X_{j\alpha} - \bar{X}_j)' \qquad i = 1,2 ; \quad j = 1,2 .$$

and

$$D_{0i} = [N/(N+1)](X_0 - \bar{X}_i)(X_0 - \bar{X}_i)' \qquad i = 1,2 .$$

Let

$$S = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}.$$

Maximizing $L_1(\mu_1, \mu_2, \Gamma)$ with respect to μ_1 and μ_2 we get

$$L_1(\Gamma) \equiv \sup_{\mu_1,\mu_2} L_1(\mu_1, \mu_2, \Gamma)$$

$$= c_0 |\Sigma_{22\cdot 1}|^{-N/2} |\Sigma_{11}|^{-(N+1)/2}$$

(2.4)
$$\cdot \operatorname{etr}\left[-(1/2)\, \Sigma_{22\cdot 1}^{-1} (S_{22} - \beta_{21}S_{12} - S_{21}\beta_{21}' + \beta_{21}S_{11}\beta_{21}') \right]$$

$$\cdot \operatorname{etr}\left[-(1/2)\, \Sigma_{11}^{-1} (S_{11} + D_{01}) \right] .$$

Now we shall consider cases (a) through (g) for Γ.

Case (a): Γ known.

When Γ is known, an ELR rule is to accept H_1 if, and only if,

(2.5) $$((\bar{X}_1 - X_0; \Sigma_{11})) - ((\bar{X}_2 - X_0; \Sigma_{22})) < k_1$$

for some constant k_1.

Proof: We can rewrite (2.4) as

$$L_1(\Gamma) = c_0 |\Gamma|^{-N/2} \operatorname{etr}\, [- (1/2)\Gamma^{-1} S]$$

$$\cdot\ |\Sigma_{11}|^{-1/2} \operatorname{etr}\, [- (1/2)\, \Sigma_{11}^{-1} D_{01}] .$$

From symmetry, we obtain the maximum of the likelihood under H_2 by interchanging 1 and 2 as

$$L_2(\Gamma) = c_0 |\Gamma|^{-N/2} \operatorname{etr}\, [- (1/2)\Gamma^{-1} S]$$

$$\cdot\ |\Sigma_{22}|^{-1/2} \operatorname{etr}\, [- (1/2)\, \Sigma_{22}^{-1} D_{02}] .$$

Hence (2.2) reduces to (2.5).

Case (b): Γ unknown.

When Γ is unknown, an ELR rule is to accept H_1 if, and only if,

$$(2.6) \qquad \frac{|S_{11}|[1 + \operatorname{tr} S_{11}^{-1} D_{01}]^{(N+1)}}{|S_{22}|[1 + \operatorname{tr} S_{22}^{-1} D_{02}]^{(N+1)}} < k_2$$

for some constant k_2.

Proof: From (2.4) it follows that

$$\sup_{\Gamma} L_1(\Gamma) = c_1 |S_{22\cdot 1}|^{-N/2} |S_{11} + D_{01}|^{-(N+1)/2}$$

where

$$c_1 = c_0 N^{N/2} (N+1)^{(N+1)/2} \exp[-p(2N+1)/2]$$

and

$$S_{22\cdot 1} = S_{22} - S_{21} S_{11}^{-1} S_{12} .$$

We can rewrite $\sup_{\Gamma} L_1(\Gamma)$ as

$$\sup_{\Gamma} L_1(\Gamma) = c_1 |S|^{-N/2} |S_{11}|^{-1/2} [1 + \operatorname{tr} S_{11}^{-1} D_{01}]^{-(N+1)/2} .$$

Using symmetry and by interchanging 1 and 2 we get

$$\sup_{\Gamma} L_2(\Gamma) = c_1 |S|^{-N/2} |S_{22}|^{-1/2} [1 + \operatorname{tr} S_{22}^{-1} D_{02}]^{-(N+1)/2} .$$

Hence (2.2) reduces to (2.6).

Case (c): Σ_{ii} known, $i = 1,2$, Σ_{12} unknown.

In this case, an ELR rule is to accept H_1 if, and only if, (2.5) holds.

Proof: We notice that

$$\sup_{\Sigma_{12}} L_1(\Gamma) = c_0|\Sigma_{11}|^{-1/2} \operatorname{etr}[-(1/2)\Sigma_{11}^{-1} D_{01}]$$

$$\cdot \sup_{\Sigma_{12}} [|\Gamma|^{-N/2} \operatorname{etr}\{-(1/2)\Gamma^{-1} S\}]$$

and

$$\sup_{\Sigma_{12}} L_2(\Gamma) = c_0|\Sigma_{22}|^{-1/2} \operatorname{etr}[-(1/2)\Sigma_{22}^{-1} D_{02}]$$

$$\cdot \sup_{\Sigma_{12}} [|\Gamma|^{-N/2} \operatorname{etr}\{-(1/2)\Gamma^{-1} S\}] .$$

Hence (2.2) reduces to (2.5).

Case (d): $\Sigma_{11} = \Sigma_{22} = \Sigma$ unknown; $\Sigma_{12} = \Sigma_{21}$ unknown.

It is not possible to give a closed form for $\sup L_i(\mu_1, \mu_2, \Gamma)$. However, this may be evaluated iteratively from the likelihood equations given below.

It can be seen that

$$L_1(\Gamma) = \text{constant } |\tau_{11}|^{-N/2} |\tau_{22}|^{-N/2} |(\tau_{11} + \tau_{22})/4|^{-1/2}$$

$$\cdot \operatorname{etr}[-(1/2)\tau_{11}^{-1}(S_{11} + S_{22} + S_{12} + S_{21})]$$

$$\cdot \operatorname{etr}[-(1/2)\tau_{22}^{-1}(S_{11} + S_{22} - S_{12} - S_{21})]$$

$$\cdot \operatorname{etr}\,[-\,(1/2)\,\{(\tau_{11} + \tau_{22})/4\}^{-1}\, D_{01}\,]$$

where

$$\tau_{11} = 2(\Sigma + \Sigma_{12}), \quad \tau_{22} = 2(\Sigma - \Sigma_{12}) .$$

Differentiating $L_i(\Gamma)$ with respect to the elements of τ_{11} and τ_{22} we get the following likelihood equations, for $i = 1,2$

$$0 = N[2\Gamma_{ii}^{-1} - D_{\Gamma_{ii}^{-1}}] + [2(\Gamma_{11} + \Gamma_{22})^{-1} - D_{(\Gamma_{11} + \Gamma_{22})^{-1}}]$$

$$+ [D_{\Gamma_{ii}^{-1} S_{ii}^{*} \Gamma_{ii}^{-1}} - \Gamma_{ii}^{-1} S_{ii}^{*} \Gamma_{ii}]$$

$$+ [D_{(\Gamma_{11} + \Gamma_{22})^{-1} D_{0i}^{*} (\Gamma_{11} + \Gamma_{22})^{-1}}$$

$$- (\Gamma_{11} + \Gamma_{22})^{-1} D_{0i}^{*} (\Gamma_{11} + \Gamma_{22})^{-1}]$$

where

$$(2.7) \qquad \begin{aligned} S_{11}^{*} &= S_{11} + S_{22} + S_{12} + S_{21} \\ S_{22}^{*} &= S_{11} + S_{22} - S_{12} - S_{21} \end{aligned}$$

and $\qquad D_{0i}^{*} = 4\, D_{0i} \quad i = 1,2.$ D_A is a diagonal matrix and D_A and A have some diagonal elements.

Case (e): $\Sigma_{11} = \Sigma_{22} = \Sigma$ known ; $\Sigma_{12} = \rho\, \Sigma$, ρ unknown.

Since this is a special case of case (c), an ELR rule accepts H_1 if, and only if, (2.8) $\quad ((\bar{x}_1 - x_0 ; \Sigma)) - ((\bar{x}_2 - x_0 ; \Sigma)) < k_3$

for some constant k_3.

Case (f): $\Sigma_{11} = \Sigma_{22} = \Sigma$ unknown; $\Sigma_{12} = \rho\,\Sigma$, ρ known.

In this case, an ELR rule is to accept H_1 if, and only if,

(2.9) $$\frac{1 + \operatorname{tr} A^{-1} D_{01}}{1 + \operatorname{tr} A^{-1} D_{02}} < k_4$$

for some constant k_4 and where

(2.10) $\quad A = (S_{11} - \rho S_{12} - \rho S_{21} + S_{22})/(1 - \rho^2)$.

Proof: We note that in this case

$$\beta_{21} = \rho I_p$$

and

$$\Sigma_{22\cdot 1} = (1-\rho^2)\,\Sigma\ .$$

Hence (2.4) becomes

$$L_1(\Gamma) = c_0(1-\rho^2)^{-N/2}\ |\Sigma|^{-(2N+1)/2}$$

$$\cdot\ \operatorname{etr}\,[-\ (1/2)\,\Sigma^{-1}\,(A + D_{01})]\ .$$

Hence, for $i = 1,2$

$$\sup_{\Sigma} L_i = c_0 (1-\rho^2)^{-N/2} (2N+1)^{(2N+1)/2}$$

$$\cdot \exp [- p(2N+1)/2] \quad |A + D_{0i}|^{-(2N+1)/2} .$$

Hence (2.2) reduces to (2.9).

Case (g): $\Sigma_{11} = \Sigma_{22} = \Sigma$ unknown; $\Sigma_{12} = \rho \Sigma$, ρ unknown.

This is a special case of case (d) considered earlier. Here

$$\tau_{11} = 2(1+\rho) \Sigma , \quad \tau_{22} = 2(1-\rho) \Sigma .$$

However, this case gives rise to another difficulty, since τ_{11} and τ_{22} are proportional. In this case the ML estimate of Σ is, under H_i ,

$$S_{11}^*/\lambda + S_{22}^* + D_{0i}^*/(1+\lambda) \qquad i = 1,2$$

where S_{ii}^* and D_{0i}^* are given by (2.7) and

$$\lambda = (1-\rho)/(1+\rho) .$$

The likelihood equation of λ is a 3p degree polynomial in λ and is given by, for $p = 1$

$$0 = a_3 \lambda^3 + a_2\lambda^2 + a_1\lambda + a_0$$

where

$$a_3 = - (N+1) S_{22}^*$$
$$a_2 = N(S_{11}^* + D_{0i}^*) - (2N+1) S_{22}^*$$
$$a_1 = (2N+1) S_{11}^* - N(S_{22}^* + D_{0i}^*)$$

and

$$a_0 = (N+1)\, S_{11}^{*}.$$

2.2. PLR rules.

As we have done for the ELR rules, we shall compute the PLR rules for cases (a) through (g). It is very straight forward to find PLR rules except for case (g), in which case we shall consider only $p = 1$. We first note that the rule given by (2.1) is equivalent to the rule which accepts H_1 if, and only if,

$$(2.11)\quad \log(|\Sigma_{22}|/|\Sigma_{11}|) + ((X_0 - \mu_2 \,;\, \Sigma_{22})) - ((X_0 - \mu_1 \,;\, \Sigma_{11})) > c$$

for some constant c.

The ML estimates of μ_1 and μ_2 are respectively $\bar{X}_1$ and $\bar{X}_2$. Hence, when Σ_{11} and Σ_{22} are known, (which is assumed in cases (a), (c) and (e)), a PLR rule accepts H_1 if, and only if,

$$((X_0 - \bar{X}_2 \,;\, \Sigma_{22}) - ((X_0 - \bar{X}_1 \,;\, \Sigma_{11})) > c_1$$

for some constant c_1.

For case (b), a PLR rule is very similar to the ELR given by (2.6). Since the ML estimate of Σ_{ii} based only on the training sample is S_{ii}, a PLR rule is to accept H_1 if, and only if,

$$(2.12)\quad \log(|S_{22}|/|S_{11}|) + N[((X_0 - \bar{X}_2 \,;\, S_{22})) - ((X_0 - \bar{X}_1 \,;\, S_{11}))] > c_2$$

for some constant c_2.

For the case (d), the ML estimate of Σ is

$$(S_{11} + S_{22})/2N \,.$$

Hence a PLR rule accepts H_1 if, and only if,

$$(2.13) \qquad ((X_0 - \overline{X}_2;\ S_{11} + S_{22})) - ((X_0 - \overline{X}_1;\ S_{11} + S_{22})) > c_3$$

for some constant c_3.

For the case (f), the ML estimate of Σ is $A/2N$, where A is given by (2.10). Hence a PLR rule is to accept H_1 if, and only if,

$$(2.14) \qquad \mathrm{tr}[A^{-1} D_{02} - A^{-1} D_{01}] > c_4$$

for some constant c_4.

For the case (g), Kim [17] has obtained the likelihood equations and studied the roots. However, they cannot be written in a closed form except when $p = 1$. For $p = 1$, we get back the case (d), and PLR rule is the same as given in (2.13). For $p > 1$, the rule given in (2.13) may be considered as an approximate PLR.

3. DISTRIBUTIONS AND REPRESENTATIONS OF CLASSIFICATION STATISTICS.

In the sequel, we shall assume

$$(3.1) \qquad \Gamma = \begin{pmatrix} \Sigma & \rho\Sigma \\ \rho\Sigma & \Sigma \end{pmatrix} .$$

This structure of Γ will be heavily used to reduce our problem, whenever possible, to a standard problem by suitable transformations of random variables involved in a classification statistic. As we shall see later, the distributions are very complicated in structure and the calculations of percentage points are very involved.

3.1. Σ known ; ρ known or unknown.

In this case both ELR and PLR rules are based on the statistic

$$(3.2) \qquad T_0 = ((\overline{X}_2 - X_0 \,;\, \Sigma)) - ((\overline{X}_1 - X_0 \,;\, \Sigma)) .$$

Without loss of any generality we can take $\Sigma = I_p$, since Σ is known.

Define B_1 and B_2 as

$$\sqrt{2}\, B_1 = [(\overline{X}_1 - \overline{X}_2)/\sqrt{c_1} - (\overline{X}_1 + \overline{X}_2 - 2X_0)/\sqrt{c_2}\ ]$$

and

$$\sqrt{2}\, B_2 = [(\overline{X}_1 - \overline{X}_2)/\sqrt{c_1} + (\overline{X}_1 + \overline{X}_2 - 2X_0)/\sqrt{c_2}\ ]$$

where

$$(3.3) \qquad c_1 = 2(1-\rho)/N \ , \quad C_2 = 4 + 2(1+\rho)/N .$$

Under H_1, B_1 and B_2 are independently distributed according to the p-variate normal distribution with means $k_1\delta$ and $k_2\delta$, respectively, and with the same covariance matrix I_p, where

$$\delta = \mu_1 - \mu_2$$

and

$$\sqrt{2}\, k_i = 1/\sqrt{c_1} + (-1)^{1+i}/\sqrt{c_2} \qquad i = 1,2, .$$

Hence we can write

$$T_0^* \equiv 2T_0/\sqrt{c_1 c_2} = B_1' B_1 - B_2' B_2 .$$

But $B_i' B_i$ follows a noncentral chi-square distribution with p degrees of freedom, under H_1, with parameter of noncentrality

$$\lambda_i = k_i^2\, \delta'\delta \qquad i = 1,2 .$$

Thus T_0^* equals the difference of two independent noncentral chi-square random variables. The density function of T_0^*, given by $f(t_0^*)$, is well known and is given by (see [14])

$$f(t_0^*) = 2^{-p} e^{-(\lambda_1+\lambda_2)} \sum_{\gamma=0}^{\infty} \sum_{s=0}^{\infty} \left[\frac{2^{-(\gamma+s)} \lambda_1^{\gamma} \lambda_2^{s}}{\gamma!\, s!\, \Gamma(s+p/2)} \cdot (-t_0^*)^{(\gamma+s+p-2)/2} W_{-\ell,m}(-t_0^*) \right] \qquad \text{if} \quad t_0^* < 0$$

and

$$f(t_0^*) = 2^{-p} e^{-(\lambda_1+\lambda_2)} \sum_{\gamma=0}^{\infty} \sum_{s=0}^{\infty} \left[\frac{2^{-(\gamma+s)} \lambda_1^{\gamma} \lambda_2^{s}}{\gamma!\, s!\, \Gamma(\gamma+p/2)} \right.$$

$$\cdot\; t_0^{*(\gamma+s+p-2)/2} \quad W_{\ell,m}(t_0^*)\Big] \qquad \text{if } t_0^* > 0$$

where

$$\ell = (\gamma-s)/2\ , \quad m = (\gamma + s + p - 1)/2$$

and

$$W_{\ell,m}(x) = \frac{x^{m+1/2}\, e^{-x/2}}{\Gamma(m-\ell+1/2)} \int_0^\infty e^{-tx}\, t^{m-\ell-1/2}(1+t)^{m+\ell-1/2}\, dt\ .$$

3.2. Σ unknown but ρ known.

In the case when Σ is unknown but ρ is known, both EML and PLR rules are based on the statistic

$$(3.4) \qquad T_1 = ((\bar{X}_2 - X_0\ ;\ A)) - ((\bar{X}_1 - X_0\ ;\ A))$$

where A is given by (2.10). We can, however, rewrite A as

$$A = (S_{11} - \rho S_{12} - \rho S_{21} + S_{22})/(1-\rho^2)$$

$$= (S_{11} - S_{12} - S_{21} + S_{22})/2(1-\rho)$$

$$+ (S_{11} + S_{12} + S_{21} + S_{22})/2(1+\rho)\ .$$

Let

$$X_{1\alpha}^* = X_{1\alpha} - X_{2\alpha}$$

$$X_{2\alpha}^* = X_{1\alpha} + X_{2\alpha}\ .$$

Then

$$\begin{pmatrix} X_{1\alpha}^* \\ X_{2\alpha}^* \end{pmatrix} \sim N_{2p}\left[\begin{pmatrix} \mu_1 - \mu_2 \\ \mu_1 + \mu_2 \end{pmatrix},\ \Gamma^*\right]$$

where

$$\Gamma^* = \begin{pmatrix} 2(1-\rho)\Sigma & 0 \\ 0 & 2(1+\rho)\Sigma \end{pmatrix} .$$

Define, for $i = 1,2$

$$(3.5) \qquad \bar{X}_i^* = \sum_{\alpha=1}^{N} X_{i\alpha}^*/N$$

and

$$(3.6) \qquad S_{ii}^* = \sum_{\alpha=1}^{N} (X_{i\alpha}^* - \bar{X}_i^*)(X_{i\alpha}^* - \bar{X}_i^*)' .$$

Then $\bar{X}_1^*$, $\bar{X}_2^*$, S_{11}^*, S_{22}^* are mutually independently distributed and

$$(3.7) \qquad S_{11}^*/2(1-\rho) = (S_{11} + S_{22} - S_{12} - S_{21})/2(1-\rho) \sim W_p[\Sigma, N-1]$$

$$(3.8) \qquad S_{22}^*/2(1+\rho) = (S_{11} + S_{22} + S_{12} + S_{21})/2(1+\rho) \sim W_p[\Sigma, N-1] .$$

Hence

$$(3.9) \qquad A = S_{11}^*/2(1-\rho) + S_{22}^*/2(1+\rho) \sim W_p[\Sigma, 2(N-1)] .$$

We can rewrite T_1, given by (3.4), as

$$T_1 = (2X_0 - \bar{X}_2^*)' A^{-1} \bar{X}_1^* .$$

We shall find the distribution of

$$M_{12} \equiv T_1/\sqrt{c_1 c_2}$$

where c_1 and c_2 are given by (3.3).

Let

$$Z = \bar{X}_1^*/\sqrt{c_1}\ , \quad Z^* = (2X_0 - \bar{X}_2^*)/\sqrt{c_2}\ . \tag{3.10}$$

Then Z and Z^* are independently distributed and

$$Z \sim N_p\ [(\mu_1 - \mu_2)/\sqrt{c_1}\ , \Sigma\ ]$$

and under H_i, $i = 1,2$

$$Z^* \sim N_p[(-1)^{i+1}\ (\mu_1 - \mu_2)/\sqrt{c_2}\ , \Sigma\ ]\ .$$

Now

$$M_{12} = Z^{*'}\ A^{-1}\ Z\ .$$

Let

$$M_{11} = Z^{*'}\ A^{-1}\ Z^* \quad \text{and} \quad M_{22} = Z'\ A^{-1}\ Z\ .$$

Then, for $p > 1$, the joint density of M_{12}, M_{11}, and M_{22} is given by Sitgreaves [23] as, under H_1

$$f_1(m_{12},m_{11},m_{22}) = d(m_{11}m_{22}-m_{12}^2)^{(p-3)/2}$$

$$\cdot \sum_{j=0}^{\infty} d_j^* \frac{[m_{11}/c_2+2m_{12}/\sqrt{c_1c_2} + m_{22}/c_1+(1/c_1+1/c_2)(m_{11}m_{22}-m_{12}^2)]^j}{[(1+m_{11})(1+m_{22}) - m_{12}^2]^{N+j}}$$

$$= d \cdot (m_{11}m_{22} - m_{12}^2)^{(p-3)/2} \tag{3.11}$$

$$\cdot \sum_{j=0}^{\infty} d_j \frac{[c_1m_{11}+2\sqrt{c_1c_2}\ m_{12}+c_2m_{22}+ (c_1+c_2)(m_{11}m_{22}-m_{12}^2)]^j}{[(1+m_{11})(1+m_{22}) - m_{12}^2]^{N+j}}$$

$$0 \le m_{ii} < \infty,\ (m_{11}m_{22}-m_{12}^2) \ge 0$$

where

$$d = \frac{\Gamma(\frac{2N-1}{2})\, e^{-\alpha^2(1/c_1+1/c_2)/2}}{\Gamma(N-p/2)\Gamma(N-(p+1)/2)\Gamma(\frac{p-1}{2})\Gamma(1/2)}$$

and

$$d_j \equiv d_j^*(c_1c_2)^{-j}$$

(3.12)
$$= \frac{\Gamma(N+j)(\alpha^2/2c_1c_2)^j}{\Gamma(j+p/2)\ j!}$$

and

(3.13) $$\alpha = ((\mu_1 - \mu_2; \Sigma)) .$$

Substituting the values of c_1 and c_2 from (3.3) we have

(3.14) $$c_1m_{11}+2\sqrt{c_1c_2}\ m_{12} + c_2m_{22} + (c_1+c_2)(m_{11}m_{22}-m_{12}^2)$$

$$= \frac{2}{N}\ [(1-\rho)m_{11}+2\sqrt{(1-\rho)(2N+1+\rho)}\ m_{12} + (2N+1+\rho)\ m_{22}$$

$$+ 2(N+1)(m_{11}m_{22}-m_{12}^2)] .$$

We proceed to find the density of M_{12} by making transformations as was done by Sitgreaves [24] for the standard case.

Let

$$M_{11} = \frac{M_{22}^2 + M_{22} + M_{12}^2}{M_{22}(1 + M_{22})}\ T + \frac{M_{12}^2}{M_{22}} .$$

Jacobian of transformation from M_{11} to T is

$$(m_{22}^2 + m_{22} + m_{12}^2)/m_{22}(1 + m_{22}) .$$

Hence, density of M_{12}, T and M_{22} becomes, from (3.11) and (3.14)

$$f_1(m_{12},t,m_{22}) = d\left[\frac{m_{22}^2+m_{22}+m_{12}^2}{m_{22}(1+m_{22})}\,t\right]^{(p-3)/2} \cdot \frac{(m_{22}^2+m_{22}+m_{12}^2)}{m_{22}(1+m_{22})}$$

$$\cdot \sum_{j=0}^{\infty} d_j \left[\frac{(1-\rho)\{\frac{m_{22}^2+m_{22}+m_{12}^2}{m_{22}(1+m_{22})}t + \frac{m_{12}^2}{m_{22}}\} + 2\sqrt{(1-\rho)(2N+1+\rho)}\,m_{12} + (2N+1+\rho)m_{22}}{\{\frac{m_{22}^2+m_{22}+m_{12}^2}{m_{22}}(1+t)\}^N}\right.$$

$$(3.15) \qquad \left. + \frac{2(N+1)\,\frac{m_{22}^2+m_{22}+m_{12}^2}{1+m_{22}}\,t}{\{\frac{m_{22}^2+m_{22}+m_{12}^2}{m_{22}}(1+t)\}^N}\right]^j \cdot \left(\frac{2}{N}\right)^j$$

$$= d\sum_{j=0}^{\infty} d_j \frac{t^{(p-3)/2}\, m_{22}^{N-1}}{(1+m_{22})^{(p-1)/2}(m_{22}^2+m_{22}+m_{12}^2)^{N+j-(p-1)/2}(1+t)^{N+j}} \cdot (2/N)^j$$

$$\cdot \left[\frac{(m_{22}^2+m_{22}+m_{12}^2)}{1+m_{22}}\{(1-\rho)+2(N+1)m_{22}\}t + \{\sqrt{1-\rho}\,m_{12} + \sqrt{2N+1+\rho}\,m_{22}\}^2\right]^j$$

$$= d\sum_{j=0}^{\infty} d_j \frac{t^{(p-3)/2}\, m_{22}^{N-1}}{(1+m_{22})^{(p-1)/2}(m_{22}^2+m_{22}+m_{12}^2)^{N+j-(p-1)/2}(1+t)^{N+j}} \cdot (2/N)^j$$

$$\cdot \left[\sum_{k=0}^{j} \frac{j!}{(j-k)!\,k!}(\sqrt{1-\rho}\,m_{12} + \sqrt{2N+1+\rho}\,m_{22})^{2(j-k)}\{(1-\rho)+2(N-1)m_{22}\}^k\right.$$

$$\left. \cdot \{(m_{22}^2 + m_{22} + m_{12}^2)/(1 + m_{22})\}^k\, t^k\right]$$

$$= d \sum_{k=0}^{\infty} \sum_{j=0}^{\infty} \left[\frac{\Gamma(N+j+k)(\alpha^2/2)^{j+k}(2/Nc_1c_2)^{j+k}\, t^{(p-3)/2+k}}{\Gamma(p/2+j+k)\; k!\; j!\; (1+t)^{N+j+k}} \right.$$

$$\left. \cdot \frac{m_{22}^{(N-1)}(\sqrt{1-\rho}\; m_{12} + \sqrt{2N+1+\rho}\; m_{22})^{2j}[1-\rho+2(N-1)m_{22}]^k}{(1+m_{22})^{(p-1)/2+k}(m_{22}^2+m_{22}+m_{12}^2)^{N+j-(p-1)/2}} \right]$$

the range of t is from 0 to ∞.

If we examine the infinite series in (3.11) and since

$$[m_{11}(1+m_{22})-m_{12}^2][m_{22}(1+m_{11})-m_{12}^2]-m_{12}^2 = [(1+m_{11})(1+m_{22})-m_{12}^2](m_{11}m_{22}-m_{12}^2) \geq 0$$

we have

$$\frac{c_1m_{11} + 2\sqrt{c_1c_2}\; m_{12} + c_2m_{22} + (c_1+c_2)(m_{11}m_{22}-m_{12}^2)}{(1+m_{11})(1+m_{22}) - m_{12}^2}$$

$$= \frac{c_1[m_{11}(1+m_{22}) - m_{12}^2] + c_2[m_{22}(1+m_{22}) - m_{12}^2] + 2(c_1c_2)^{1/2}m_{12}}{(1+m_{11})(1+m_{22}) - m_{12}^2}$$

$$< \frac{[\{c_1[m_{11}(1+m_{22}) - m_{12}^2]\}^{1/2} + \{c_2[m_{22}(1+m_{11}) - m_{12}^2]\}^{1/2}]^2}{(1+m_{11})(1+m_{22}) - m_{12}^2}$$

$$< (\sqrt{c_1} + \sqrt{c_2})^2 .$$

Hence

$$f_1(m_{12},m_{11},m_{22}) < d \sum_{j=0}^{\infty} \frac{\Gamma(N+j)(\alpha^2/2c_1c_2)^j(\sqrt{c_1} + \sqrt{c_2})^{2j}}{\Gamma(j+p/2)\; j!}$$

$$< d \sum_{j=0}^{\infty} \frac{\Gamma(N)}{\Gamma(p/2)} \frac{\{(\alpha^2/2c_1c_2)(\sqrt{c_1}+\sqrt{c_2})^2\}^j}{j!}$$

$$= d \frac{\Gamma(N)}{\Gamma(p/2)} \exp [(\alpha^2/2\, c_1\, c_2)(\sqrt{c_1}+\sqrt{c_2})^2] .$$

Hence, the infinite series in (2.11) is uniformly convergent. So we can integrate the transformed density (3.15) with respect to t term by term, and we get

$$f_1(m_{12}, m_{22}) = d \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \frac{\alpha^{2(j+k)}\Gamma(p/2-1/2+k)\Gamma(N+j-p/2+1/2)}{\Gamma(p/2+j+k)\, k!\, j!} (N\, c_1c_2)^{-(j+k)}$$

(3.16)

$$\cdot \frac{m_{22}^{N-1}[\sqrt{1-\rho}\, m_{12} + \sqrt{2N+1+\rho}\, m_{22}]^{2j}[1-\rho+2(N+1)m_{22}]^k}{(1+m_{22})^{(p-1)/2+k}(m_{22}^2+m_{22}+m_{12}^2)^{N+j-(p-1)/2}}$$

Now

$$[\sqrt{1-\rho}\, m_{12} + \sqrt{2N+1+\rho}\, m_{22}]^{2j}$$

$$= [(1-\rho)m_{12}^2 + (2N+1+\rho)\, m_{22}^2 + 2\sqrt{(1-\rho)(2N+1+\rho)}\, m_{12}\, m_{22}]^j$$

$$= \sum_{s=0}^{j} \sum_{r=0}^{s} \frac{j!}{r!(s-r)!(j-s)!} 2^{s-r}(1-\rho)^{r+(s-r)/2}(2N+1+\rho)^{(j-s)+(s-r)/2}$$

$$\cdot\, m_{12}^{2r+(s-r)}\, m_{22}^{2(j-s)+(s-r)}$$

and

$$[1-\rho+2(N+1)m_{22}]^k = \sum_{m=0}^{k} \frac{k!}{m!(k-m)!} (1-\rho)^{k-m}\, 2^m (N+1)^m\, m_{22}^m .$$

Hence (3.16) can be written as

$$f_1(m_{12}, m_{22}) =$$

$$d \sum_{j=0}^{\infty} \sum_{s=0}^{j} \sum_{r=0}^{s} \sum_{k=0}^{\infty} \sum_{m=0}^{k} \frac{\alpha^{2(j+k)} \Gamma(p/2-1/2+k) \Gamma(N+j-(p-1)/2)(Nc_1c_2)^{-(j+k)}}{\Gamma(p/2+j+k) r!\,(s-r)!\,(j-s)!\,(k-m)!}$$

$$\cdot\, m_{22}^{N-1} \quad 2^{s-r} (1-\rho)^{r+(s-r)/2} (2N+1+\rho)^{(j-s)+(s-r)/2} \quad m_{12}^{2r+(s-r)}$$

$$\cdot (1-\rho)^{k-m}\, 2^m (N+1)^m\, m_{22}^{m+2(j-s)+(s-r)}$$

$$\cdot\, [(1+m_{22})^{(p-1)/2+k} (m_{22}^2 + m_{22} + m_{12}^2)^{N+j-(p-1)/2}]^{-1}$$

$$= d \sum_{j=0}^{\infty} \sum_{s=0}^{\infty} \sum_{r=0}^{\infty} \sum_{k=0}^{\infty} \sum_{m=0}^{\infty} \frac{\alpha^{2(j+k+r+s+m)} \Gamma(k+(p-1)/2+m)}{\Gamma(p/2+j+r+s+k+m)\; r!\; s!\; j!\; k!}$$

$$\cdot\, \Gamma(N+j+r+s-(p-1)/2)(N\, c_1c_2)^{j+s+r+k+m}$$

$$\cdot\, \frac{(N+1)^m (1-\rho)^{r+k+s/2}\, 2^{s+m} (2N+1+\rho)^{j+s/2}\, m_{22}^{N-1+2j+s+m}\, m_{12}^{2r+s}}{(1+m_{22})^{(p-1)/2+m+k} (m_{22}^2+m_{22}+m_{12}^2)^{N+j+r+s-(p-1)/2}} .$$

Integrating m_{22} we get

$$(3.17) \qquad f_1(m_{12}) = d \sum_{j=0}^{\infty} \sum_{s=0}^{\infty} \sum_{r=0}^{\infty} \sum_{k=0}^{\infty} \sum_{m=0}^{\infty} \alpha^{2(j+k+\gamma+\alpha+m)}$$

$$\cdot \frac{\Gamma(k+(p-1)/2+m)\Gamma(N+j+r+s-(p-1)/2)}{\Gamma(p/2+j+r+s+k+m)\ r!\ s!\ j!\ k!}$$

$$\cdot (N+1)^m (1-\rho)^{r+k+s/2}\, 2^{s+m} (2N+1+\rho)^{j+s/2} \quad (N\, c_1 c_2)^{j+s+r+k+m}$$

$$\cdot F(m_{12},\ 2r+s,\ N-1+2j+s+m,\ (p-1)/2+m+k,\ N+j+r+s-(p-1)/2)$$

where, for nonnegative a_1, a_2, a_3, a_4 and $-\infty < x < \infty$,

$$(3.18)\quad F(x,\ a_1,a_2,a_3,a_4) = x^{a_1} \int_0^\infty \frac{y^{a_2}}{(1+y)^{a_3}(y^2+y+x^2)^{a_4}}\, dy\ .$$

Let $y = z/(1-z)$. Then Jacobian is $(1-z)^{-2}$. Therefore,

$$F(x,a_1,a_2,a_3,a_4) = x^{a_1} \int_0^1 \frac{z^{a_2}(1-z)^{2a_4+a_3-a_2-2}}{[z + (1-z)^2 x^2]^{a_4}}\, dz\ .$$

To perform the integration we follow Sitgreaves [24]. Let $x^2 < 1$. Then

$$z + (1-z)^2 x^2 = 1 - (1-z)(1-x^2+x^2 z)$$

and

$$(1-z)(1 - x^2 + x^2 z) \leq 1\ .$$

Hence

$$[z + (1-z)^2 x^2]^{-a_4} = \sum_{a=0}^{\infty} \frac{(a_4 + a)!}{a!\ (a_4)!} (1-z)^a (1 - x^2 + x^2 z)^a$$

$$= \sum_{r=0}^{\infty} \frac{(a_4 + a)!}{a!\ (a_4)!} (1-z)^a \sum_{b=0}^{a} \frac{a!}{b!\ (a-b)!} (1-x^2)^{a-b} (x^2 z)^b$$

$$= \sum_{b=0}^{\infty} \sum_{a=0}^{\infty} \frac{(a_4 + a + b)!}{a!\ b!\ a_4!} (1-z)^{a+b} (1-x^2)^a\ x^{2b}\ z^b\ .$$

Since the series expansion is uniformly convergent for $|x| \le 1$, we can integrate term by term and, for $x^2 \le 1$

$$F(x,a_1,a_2,a_3,a_4) = \sum_{b=0}^{\infty} \sum_{a=0}^{\infty} \frac{(a_4 + a + b)!}{a!\ b!\ a_4!} (1-x^2)^a\ x^{2b+a_1}$$

$$\cdot \frac{\Gamma(b + a_2 + 1)\Gamma(a_3 + 2a_4 - a_2 - 1 + a + b)}{\Gamma(a_3 + 2a_4 + a + 2b)}\ .$$

When $x^2 > 1$,

$$z + x^2(1-z)^2 = x^2[1-z(1-z + \frac{x^2-1}{x^2})]$$

and

$$z[1 - z + (x^2 - 1)/x^2] \le 1\ .$$

Hence

$$[z + x^2(1-z)^2]^{-a_4}$$

$$= x^{-2a_4} \sum_{a=0}^{\infty} \frac{(a_4 + a)!}{a!\ a_4!}\ z^a(1 - z + \frac{x^2 - 1}{x^2})^a$$

$$= x^{-2a_4} \sum_{a=0}^{\infty} \frac{(a_4 + a)!}{a!\ a_4!}\ z^a \sum_{b=0}^{a} (1-z)^b \left(\frac{x^2 - 1}{x^2}\right)^{a-b} \cdot \frac{a!}{b!\ (a-b)!}$$

$$= x^{-2a_4} \sum_{b=0}^{\infty} \sum_{a=0}^{\infty} \frac{(a_4 + a + b)!}{a_4!\ a!\ b!} \left(\frac{x^2 - 1}{x^2}\right)^a z^{a+b}(1-z)^b\ .$$

Hence, for $x^2 > 1$

$$F(x,a_1,a_2,a_3,a_4) = \sum_{b=0}^{\infty} \sum_{a=0}^{\infty} \frac{(a_4 + a + b)!}{a!\, b!\, a_4!} x^{a_1-2a_4-2a} (x^2 - 1)^a$$

$$\cdot \frac{\Gamma(a_2+a+b+1)\Gamma(a_3+2a_4-a_2+b-1)}{\Gamma(a_3+2a_4+2b+a)} .$$

Now consider $p = 1$. Under H_1, joint density of Z^*, Z and A is

$$(3.19) \quad f_1(z^*,z,a) = \frac{e^{-a/2}\, a^{n-1}\, e^{-\frac{1}{2}(z^*-\delta/\sqrt{c_2})^2-\frac{1}{2}(z-\delta/\sqrt{c_1})^2}}{\Gamma(n)\, 2^n\, 2\pi}$$

$$= c \cdot e^{-a/2}\, a^{n-1}\, e^{-(z^{*2}+z^2)/2+\delta(z^*/\sqrt{c_2}+z/\sqrt{c_1})}$$

where

$$n = 2(N-1) \quad \text{and} \quad c = \frac{e^{-\delta^2(1/c_1+1/c_2)/2}}{\Gamma(n)\, 2^{n+1}\, \pi} .$$

Let

$$U_1 = Z^*/\sqrt{A} , \quad U_2 = Z/\sqrt{A} .$$

Jacobian of transformation is a. Hence the joint density of U_1, U_2, and A is

$$(3.20) \quad f_1(u_1,u_2,a) = c\, e^{-a(1+u_1^2+u_2^2)/2}\, a^n\, e^{(u_1/\sqrt{c_1} + u_2/\sqrt{c_2})\delta\sqrt{a}}$$

$$= c\, e^{-a(1+u_1^2+u_2^2)/2}\, a^n \sum_{\ell=0}^{\infty} \frac{\delta^\ell (u_1/\sqrt{c_1} + u_2/\sqrt{c_2})^\ell a^{\ell/2}}{\ell\,!} .$$

If we examine (3.19) we see that

$$f_1(z^*, z, a) < (2^{n+1}\pi)^{-1} e^{-(n-1)/2} (n-1)^{n-1}$$

for all z^*, z and a. Hence we can integrate a in (3.20) term by term to get the joint density of U_1 and U_2 as

$$f_1(u_1,u_2) = c \sum_{\ell=0}^{\infty} \frac{\delta^\ell (u_1/\sqrt{c_1} + u_2/\sqrt{c_2})^\ell \, \Gamma(n+\ell/2+1) 2^{n+\ell/2+1}}{\ell! \quad (1+u_1^2+u_2^2)^{n+\ell/2+1}}$$

$$= c \sum_{\ell=0}^{\infty} \frac{\delta^\ell}{\ell!} \sum_{r=0}^{\ell} \frac{\ell! \, u_1^r u_2^{\ell-r} \, \Gamma(n+\ell/2+1) \, 2^{n+\ell/2+1}}{r!(\ell-r)! \, c_1^{r/2} \, c_2^{(\ell-r)/2} (1+u_1^2+u_2^2)^{n+\ell/2+1}}$$

$$= \sum_{r=0}^{\infty} \sum_{\ell=0}^{\infty} c_{r\ell} \, u_2^\ell \, (1+u_1^2+u_2^2)^{-(n+r/2+\ell/2+1)} u_1^r$$

where

$$c_{r\ell} = c \, \frac{\delta^{\ell+r} \, \Gamma(n+r/2+\ell/2+1) \, 2^{n+r/2+\ell/2+1}}{r! \, \ell! \, c_1^{r/2} \, c_2^{\ell/2}} \, .$$

Now $\qquad M_{12} = U_1U_2 \, .$

For $t < 0$,

$$P[M_{12} < t] = P[U_1U_2 < t]$$

$$= \int_{-\infty}^{0}\int_{t/u_1}^{\infty} f_1(u_1, u_2)\, du_2 du_1 + \int_0^\infty \int_{-\infty}^{t/u_1} f_1(u_1,u_2) du_2 du_1$$

$$= \int_{-\infty}^{0}\int_{-\infty}^{t} f_1(u_1, m_{12}/u_1)\, u_1^{-1}\, dm_{12} du_1$$

$$+ \int_0^\infty \int_{-\infty}^{t} f_1(u_1, m_{12}/u_1)\, u_1^{-1}\, dm_{12} du_1$$

$$= \int_{-\infty}^{t} \int_{-\infty}^{\infty} f_1(u_1, m_{12}/u_1)\, u_1^{-1}\, du_1\, dm_{12} .$$

Similarly for $t > 0$

$$P[M_{12} > t] = \int_{t}^{\infty} \int_{-\infty}^{\infty} f_1(u_1, m_{12}/u_1)\, u_1^{-1}\, du_1\, dm_{12} .$$

For $t = 0$

$$P[M_{12} < 0] = P[U_1 U_2 < 0]$$

$$= \int_{-\infty}^{0} \int_{0}^{\infty} f_1(u_1, u_2)\, du_2 du_1 + \int_{0}^{\infty} \int_{-\infty}^{0} f_1(u_1, u_2) du_2 du_1$$

$$= \int_{-\infty}^{0} \int_{-\infty}^{\infty} f_1(u_1, m_{12}/u_1)\, u_1^{-1}\, du_1\, dm_{12}.$$

Hence the density of M_{12} is

$$f_1(m_{12}) = \int_{-\infty}^{\infty} f_1(u_1, m_{12}/u_1)\, u_1^{-1} du_1$$

$$(3.21) \qquad = \sum_{r=0}^{\infty} \sum_{\ell=0}^{\infty} c_{r\ell}\, m_{12}^{\ell} \int_{-\infty}^{\infty} u_1^{r-\ell-1} (1 + u_1^2 + m_{12}^2/u_1^2)^{-s}\, du_1$$

$$= \begin{cases} \displaystyle\sum_{r=0}^{\infty} \sum_{\ell=0}^{\infty} c_{r\ell}\, m_{12}^{\ell}\; 2\int_{0}^{\infty} \omega^{(r-\ell)/2-1} (1+\omega+m_{12}^2/\omega)^{-s}\, d\omega & \text{if } r-\ell \text{ is even} \\ 0 \qquad \text{if } r-\ell \text{ is odd} . \end{cases}$$

where $s = n + (r+\ell)/2 + 1.$

To see (3.21), we note that

$$\int_0^\infty \omega^{(r-\ell)/2-1}(1 + \omega + m_{12}^2/\omega)^{-s}\, d\omega$$

$$= \int_0^1 z^{n+r}\ (1-z)^{n+\ell}[z + m_{12}^2(1-z)^2]^{-s}\, dz$$

$$< \int_0^1 [z + m_{12}^2\ (1-z)^2]^{-s}\, dz < g(m_{12}^2)$$

where

$$g(m_{12}^2) = \sup_{0<z<1} [z + m_{12}^2(1-z)^2]^{-s} .$$

But

$$m_{12}^\ell \int_0^\infty \omega^{(r-\ell)/2-1}(1 + \omega + m_{12}^2/\omega)^{-s}\, d\omega$$

$$= m_{12}^\ell \int_0^\infty \omega^{n+r}(\omega^2 + \omega + m_{12}^2)^{-s}\, d\omega$$

$$= F(m_{12},\ \ell,\ n+r,\ 0,\ n + (r+\ell)/2 + 1)$$

where F function is given by (3.8). Hence the density of M_{12} under H_1, when $p = 1$, is

$$(3.22)\quad f_1(m_{12}) = \sum_{r=0}^\infty \sum_{\ell=0}^\infty c_{r\ell}\ [1 + (-1)^{r-\ell+1}]F(m_{12},\ \ell,\ n+r,\ 0,\ n+(r+\ell)/2+1) .$$

3.3. <u>Σ unknown, δ unknown.</u>

In the case when both Σ and δ are unknown, the PLR rule is based on the statistic, when $p = 1$,

$$(3.23) \qquad T_2 = ((X_0 - \bar{X}_1;\ S_{11} + S_{22})) - ((X_0 - \bar{X}_2;\ S_{11} + S_{22}))$$

Since the same PLR rule (2.13) is also suggested as an approximate PLR rule for $p > 1$, we shall determine the density of T_2 for all values of p. We can rewrite T_2 as

$$T_2 = \sqrt{c_1 c_2}\ \ Z^{*'}(S_{11} + S_{22})^{-1}\ Z$$

where Z^* and Z are given by (3.10) and c_1 and c_2 are given by (3.3).

Let

$$V \equiv S_{11} + S_{22} = (S_{11}^* + S_{22}^*)/2$$

where S_{ii}^* is given by (3.6). Hence it follows from (3.7) and (3.8) that

$$V = (1-\rho)V_1 + (1+\rho)V_2$$

where

$$V_1 = S_{11}^*/2(1-\rho) \quad \text{and} \quad V_2 = S_{22}^*/2(1+\rho)$$

are independently distributed as $W_p[\,\Sigma, N-1]$.

We shall find the density of

$$T_2^* \equiv 2\ T_2/\sqrt{c_1 c_2} = Z^{*'}[\lambda V_1 + (1-\lambda)V_2]^{-1}\ Z$$

where

$$\lambda = (1-\rho)/2\ .$$

We shall first obtain density of V for $p = 1$. Without loss of generality, we can take $\lambda < 1/2$ $(\lambda = 1/2 \Leftrightarrow \rho = 0$ and then $2V$ has chi-square distribution with $2(N-1)$ degrees of freedom and hence the distribution follows from the results of the last section).

The joint density of V_1 and V_2 is

$$(3.24) \qquad f(v_1,v_2) = e^{-(v_1+v_2)/2}\,(v_1v_2)^{n/2-1}/\Gamma^2(n/2)2^n \qquad v_1 > 0,\ v_2 > 0$$

where

$$n = 2(N-1)\ .$$

Let $U_1 = \lambda V_1$, $U_2 = (1-\lambda)V_2$. Jacobian of transformation is $1/\lambda(1-\lambda)$. Hence the joint density of U_1 and U_2 is

$$f(u_1,\ u_2) = \frac{e^{-[u_1/\lambda+u_2/(1-\lambda)]^2}(u_1u_2)^{n/2-1}}{\Gamma^2(n/2)\ 2^n\ [\lambda(1-\lambda)]^{n/2}} \qquad u_1 > 0,\ u_2 > 0\ .$$

Let $U = U_2$, $S^* = U_1 + U_2$. Jacobian is unity. The joint density of U and S^* is

$$f(u,\ s^*) = \frac{e^{-s^*/2\lambda}\ e^{-\frac{u}{2}(\frac{1}{1-\lambda}-\frac{1}{\lambda})}[u(s^*-u)]^{n/2-1}}{\Gamma^2(n/2)\ 2^n[\lambda(1-\lambda)]^{n/2}} \qquad 0 < u < s^*\ .$$

Hence the density of S^* is

$$f(s^*) = \frac{e^{-\ s^*/2\lambda}}{\Gamma^2(n/2)\ 2^n[\lambda(1-\lambda)]^{n/2}}$$

$$\cdot \int_0^{s^*} e^{-\frac{u}{2}[\frac{1}{1-\lambda}-\frac{1}{\lambda}]}\ u^{\frac{n}{2}-1}(s^*-u)^{n/2-1}\,du \qquad s^* > 0\ .$$

Put $v = u/s^*$, $du = s^*\,dv$. Then

$$(3.25) \qquad f(s^*) = \frac{e^{-s^*/2\lambda}\ s^{*n/2}}{\Gamma^2(n/2)2^n[\lambda(1-\lambda)]^{n/2}} \int_0^1 e^{-s^*v(\frac{1}{1-\lambda}-\frac{1}{\lambda})/2}[v(s^*-vs^*)]^{n/2-1}\,dv$$

$$= \frac{e^{-s^*/2\lambda} \; s^{*\,n-1}}{\Gamma^2(n/2)2^n[\lambda(1-\lambda)]^{n/2}} \int_0^\infty \sum_{j=0}^\infty \frac{[\frac{s^*}{2}(\frac{1}{\lambda} - \frac{1}{1-\lambda})]^j}{j!} [v(1-v)]^{n/2-1}\, dv \, .$$

We can perform integration term by term since from (3.24) we see that

$$f(v_1,v_2) < e^{-(n-2)/2} (n/2-1)^{n-2} / \Gamma^2(n/2)2^n \, .$$

Hence

$$f(s^*) = \sum_{j=0}^\infty C_j \frac{e^{-s^*/2\lambda} \; s^{*\,n+j-1}}{2^{n+j} \; \lambda^{n+j} \; \Gamma(n+j) \; j!}$$

where

$$(3.26) \qquad C_j = \frac{(1-2\lambda)^j \; \Gamma(n/2+j) \; \lambda^{n/2}}{(1-\lambda)^{n/2+j} \; \Gamma(n/2)} \, .$$

Note that

$$\sum_{j=0}^\infty C_j = (\frac{\lambda}{1-\lambda})^{n/2} \sum_{j=0}^\infty \frac{\Gamma(n/2+j)[(1-2\lambda)/(1-\lambda)]^j}{j! \; \Gamma(n/2)}$$

$$= (\frac{\lambda}{1-\lambda})^{n/2} \quad [1 - \frac{1-2\lambda}{1-\lambda}]^{-n/2} = 1$$

since

$$\frac{1-2\lambda}{1-\lambda} < 1 \quad \text{as} \quad 0 < \lambda < 1/2 \, .$$

Hence the density of $S = S^*/\lambda$ is

$$f(s) = \sum C_j \; e^{-s/2} \; s^{n+j-1}/2^{n+j} \; \Gamma(n+j) \qquad s > 0 \, .$$

So $V_1 + \frac{(1-\lambda)}{\lambda} V_2$ is a chi-square random variable with $2(N + \tau)$

degrees of freedom, where τ is an integer valued random variable with

$$(3.27) \qquad P[\tau = j] = C_j = \frac{(1-2\lambda)^j \lambda^{n/2} \Gamma(n/2+j)}{\Gamma(n/2)(1-\lambda)^{n/2+j}} \qquad j = 0,1,2,\ldots .$$

Hence the density of T_2^* is, under H_1

$$(3.28) \qquad f_1(t_2^*) = \sum_{j=0}^{\infty} C_j \, f_1(t_2^*/\lambda,\ 2n + 2j) \cdot \frac{1}{\lambda}$$

where $f_1(x,m)$ is $f_1(x)$ given by (3.22) with n replaced by m.

For $p > 1$, we note that for every $p \times 1$ vector ℓ such that $\ell'\ell = 1$,

$$\ell'(V_1 + \frac{1-\lambda}{\lambda} V_2)\ell$$

has a chi-square distribution with $2(n + \tau)$ degrees of freedom, hence

$$V_1 + \frac{(1-\lambda)}{\lambda} V_2 \sim W_p[\Sigma, 2(n + \tau)]$$

where probability function of τ is given by (3.27).

Hence, for $p > 1$, density of $\lambda^{-p} T_2^*$ is given by

$$(3.29) \qquad \Sigma C_j \, f_1^* \, (\lambda^{-p} \, t_2^* \, ; \, 2n + 2j)$$

where $f_1^*(m_{12}; m)$ is $f_1(m_{12})$ given by (3.17) with n replaced by m.

4. ASYMPTOTIC EXPANSIONS.

4.0. Introduction.

We have seen in Section 3 that the exact distributions of the statistics are very complicated for the calculation of percentage points and hence they seem to have very little practical value. In this section we shall find the asymptotic expansion of the distribution function of the statistic W, given by

$$(4.1) \qquad W = (\overline{X}_1 - \overline{X}_2)' B^{-1} (2X_0 - \overline{X}_1 - \overline{X}_2)$$

where $\overline{X}_1$ and $\overline{X}_2$ are defined in (2.3) and B is a consistent estimate of Σ. We shall assume B to be independent of $\overline{X}_1$, $\overline{X}_2$ and X_0. In this sequel we shall consider

$$(4.2) \qquad \Gamma = \begin{pmatrix} \Sigma & \Sigma_{12} \\ \Sigma_{12} & \Sigma \end{pmatrix}$$

and also the special case when $\Sigma_{12} = \rho\Sigma$. As the sample size N tends to infinity, the distribution of W tends to $N(\alpha, 4\alpha)$ or $N(-\alpha, 4\alpha)$ according as H_1 or H_2 obtains, where α is the Mahalanobis distance between two populations and is given by (3.13).

In the standard equal sample classification case, B is taken as $(S_{11} + S_{22})/(2N - 2)$ where S_{11} and S_{22} are given by (2.3). Let W_α^* and W_a^* be defined as

$$(4.3) \qquad W_\alpha^* = (W - \alpha)(4\alpha)^{-1/2}$$

and

$$(4.4) \qquad W_a^* = (W - a_N)(4a_N)^{-1/2}$$

where a_N is a consistent estimate of α and is given by

$$(4.5) \qquad a_N = ((\overline{X}_1 - \overline{X}_2;\ B)) .$$

Note that, under H_1, both W_α^* and W_a^* are standard normal variates as $N \to \infty$. In the standard equal sample case, i.e., $\Sigma_{12} = 0$ and $B = (S_{11} + S_{22})/(2N - 2)$, Bowker and Sitgreaves [7] and Okamoto [19] have derived the asymptotic expansion of the distribution function of W_α^* to the order of N^{-3}; Anderson [5] has derived the asymptotic expansion of the distribution function of W_a^* to the order of N^{-2}. Anderson [4] has compared these two asymptotic expansions to the order of N^{-2}, in relation to the probability of misclassification (PMC) and their uses.

Following Anderson [5], we shall obtain expressions, to the order of N^{-2}, of the asymptotic distributions of W_α^* and W_a^* with B taken as Σ or $(B_1 + B_2)/2$ according as Σ is known or unknown, where

$$(4.6) \qquad (N-1)\, B_i \sim W_p[\,\Sigma,\ N-1] \qquad i = 1,2$$

and B_1 and B_2 are not necessarily independent. We shall assume $\alpha > 0$ and first obtain asymptotic expansions under H_1. Then we shall consider the following three special cases:

Case (a): When Σ is known, we shall take

$$(4.7) \qquad B = \Sigma .$$

In this case the W statistic reduces to $2T_0$, where T_0 is given by (3.2).

Case (b): When Σ is unknown and $\Sigma_{12} = \rho\Sigma$ with known ρ, we shall take

$$(4.8) \qquad (N-1)B_1 = S_{11}^*/2(1-\rho) \ ; \ (N-1)B_2 = S_{22}^*/2(1+\rho)$$

where S_{11}^* and S_{22}^* are given by (2.7). In this case B_1 and B_2 are independent and the W statistic reduces to $2\,T_1$, where T_1 is given by (3.4).

Case (c): When both Σ and Σ_{12} are unknown, we shall take

$$(4.9) \qquad (N-1)B_i = S_{ii} \ , \qquad i = 1,2 \ .$$

In this case the W statistic reduces to $2\,T_2$, where T_2 is given by (3.23). Hence, we can obtain the asymptotic expansions of the distribution functions of T_0, T_1 and T_2 as special cases, from the asymptotic expansion of the distribution functions of W. We shall also study the PMC of the rules associated with T_0, T_1, and T_2, given by (2.8), (2.13) and (2.14) respectively.

Without loss of any generality we can take

$$(4.10) \qquad \Sigma \equiv I_p, \ \ \mu_1 \equiv 0, \ \ -\mu_2 \equiv \delta \ , \ \ \Sigma_{12} \equiv D$$

where D is a diagonal matrix with ρ_i (cannonical correlation) as the $i^{\underline{th}}$ diagonal entry and $|\rho_i| < 1$ for all $i = 1,2,\ldots,p$.

This can be easily seen by making the following transformations

$$
(4.11)\qquad
\begin{aligned}
&\bar{X}_i \to L(\bar{X}_i - \mu_1) \qquad i = 1, 2\\
&X_0 \to L(X_0 - \mu_1)\\
&B \to LBL'
\end{aligned}
$$

where L is a $p \times p$ nonsingular matrix such that (see e.g. Rao [22])

$$
(4.12)\qquad L \Sigma L' = I_p, \quad L \Sigma_{12} L' = D
$$

and D is a diagonal matrix. Since diagonal elements of D are the latent roots of $\Sigma^{-1} \Sigma_{12}$ and the latent roots of $\Sigma - \Sigma_{12} \Sigma^{-1} \Sigma_{12}$ are positive and the same as the latent roots of $I_p - L \Sigma_{12} \Sigma^{-1} \Sigma_{12} L = I_p - D^2$, every diagonal entry of D is less than unity in absolute value. Under the transformations (4.11), $\delta = L(\mu_1 - \mu_2)$ and hence $\Delta^2 \equiv \alpha = \delta'\delta = ((\mu_1 - \mu_2; \Sigma))$.

Hence in view of (4.11) and the fact that both W_α^* and W_a^* are invariant under (4.11), we shall assume (4.10) and find the asymptotic expansions under H_1 first.

4.1. Asymptotic expansion of the distribution function of W_a^*.

The distribution function of W_a^* is given by

$$
P[W_a^* \le u] = P[U_0 \le (u\, a_N^{\frac{1}{2}} + G_{1N})\, G_{2N}^{-\frac{1}{2}}\,]
$$

where

$$
(4.12)\qquad G_{1N} = (\bar{X}_1 - \bar{X}_2)'\, B^{-1}\, \bar{X}_1
$$

$$
(4.13)\qquad G_{2N} = ((\bar{X}_1 - \bar{X}_2; B^2))
$$

and

$$(4.14)\qquad U_0 = (\bar{X}_1 - \bar{X}_2)'\, B^{-1}\, X_0\, G_{2N}^{-\frac{1}{2}} .$$

Now, conditional distribution of U_0, given $\bar{X}_1$, $\bar{X}_2$ and B, is standard normal. Hence

$$P[W_a^* \le u] = E\, \Phi[(u\, a_N^{\frac{1}{2}} + G_{1N})G_{2N}^{-\frac{1}{2}}]$$

where expectation is taken over the joint density of $\bar{X}_1$, $\bar{X}_2$ and B. Let $n = 2(N-1)$ and define Y, Z and V by

$$(4.15)\qquad \bar{X}_1 - \bar{X}_2 = \delta + Y/\sqrt{n}$$

$$(4.16)\qquad \bar{X}_1 = Z/\sqrt{n}$$

and

$$(4.17)\qquad B = I_p + V/\sqrt{n} .$$

Then

$$\begin{pmatrix} Y \\ Z \end{pmatrix} \sim N_{2p}\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix},\ \begin{pmatrix} 2(n/N)(I_p-D) & (n/N)(I_p-D) \\ (n/N)(I_p-D) & (n/N)I_p \end{pmatrix}\right]$$

and

$$a_N = ((\delta + Y/\sqrt{n}\ ;\ I_p + V/\sqrt{n}))$$

$$G_{1N} = (\delta + Y/\sqrt{n})'\ (I_p + V/\sqrt{n})^{-1}\ Z/\sqrt{n}$$

$$G_{2N} = ((\delta + Y/\sqrt{n}\ ;\ (I_p + V/\sqrt{n})^2)) .$$

Let $Y' = (Y_1,\dots,Y_p)$, $Z' = (Z_1,\dots,Z_p)$ and the element on the

$i^{\underline{th}}$ row and $j^{\underline{th}}$ column of V be V_{ij}. Let J_n be the subset of the sample space defined as

$$(4.18)\quad J_n = \{|Y_i| < 4(\log n)^{\frac{1}{2}},\ \ |Z_i| < 2(\log n)^{\frac{1}{2}},$$

$$|V_{ij}| < 2\log n\ ;\ i,j = 1,2,\ldots,p\}\ .$$

Lemma 4.1.1. $P(J_n) = 1 - o(n^{-2})$.

Proof of the Lemma is given in the Appendix. Now

$$(4.19)\quad |E\ \Phi(\{u\ a_N^{\frac{1}{2}} + G_{1N}\}\ G_{2N}^{-\frac{1}{2}})\ \chi(J_n) - E\ \Phi(\{u\ a_N^{\frac{1}{2}} + G_{1N}\}\ G_{2N}^{-\frac{1}{2}})|$$

$$= |E\ \Phi(\{u\ a_N^{\frac{1}{2}} + G_{1N}\}\ G_{2N}^{-\frac{1}{2}})\ \chi(J_n^c)|$$

$$\leq E\ \chi(J_n^c) = o(n^{-2})$$

where $\chi(A)$ stands for the indicator set function of the set A. So, to the order of n^{-2}, it is enough to consider

$$E\ \Phi(\{u\ a_N^{\frac{1}{2}} + G_{1N}\}\ G_{2N}^{-\frac{1}{2}})\ \chi(J_n)\ .$$

Using the identity

$$(I_p + n^{-\frac{1}{2}}\,V)^{-1} = I_p - n^{-\frac{1}{2}}\,V + n^{-1}\,V^2 - n^{-3/2}\,V^3$$

$$+ n^{-2}\,V^4 - n^{-5/2}\,V^5\,(I_p + n^{-\frac{1}{2}}\,V)^{-1}$$

we have, for Y and V in J_n and for large n

$$a_N = \delta'\delta + n^{-\frac{1}{2}}(2\delta'Y-\delta'V\delta) + n^{-1}(\delta'V^2\delta + Y'Y-2\delta'VY) + R_{1n}$$

where R_{1n} is a sum of two terms, given by (i) a homogeneous polynomial of degree 3 in Y and V times $n^{-3/2}$, and (ii) a term $O(\log^4 n/n^2)$. Hence, for sufficiently large n and Y, V in J_n

$$a_N^{\frac{1}{2}} = \Delta + (2\Delta\sqrt{n})^{-1}(2\delta'Y - \delta'V\delta)$$

$$+ n^{-1}[(\delta'V^2 \delta + Y'Y - 2\delta'VY)/2\Delta$$

$$- (2\delta'Y - \delta'V\delta)^2/8\Delta^3] + R_{2n}$$

where R_{2n} is defined the same way as R_{1n}, and $\Delta^2 = \alpha = \delta'\delta$. Also, we have

$$G_{1N} = n^{-\frac{1}{2}} \delta'Z + (Y'Z - \delta'VZ)n^{-1} + R_{3n}$$

where R_{3n} is the sum of three terms given by (i) a homogeneous polynomial of degree 3 in Y, Z and V multiplied by $n^{-3/2}$, (ii) a homogeneous polynomial of degree 4 in Y, Z and V multiplied by n^{-2}, and (iii) a term

$$O((\log n)^5/n^{5/2}) .$$

Now, using the identity

$$(I_p + n^{-\frac{1}{2}} V)^{-2} = I_p - 2n^{-\frac{1}{2}}V + 3n^{-1} V^2 - 4n^{-3/2} V^3$$

$$+ 5n^{-2} V^4 - n^{-5/2}(6V^5 + 5n^{-\frac{1}{2}} V^6)(I_p + n^{-\frac{1}{2}} V)^{-2}$$

we have

$$G_{2N} = \Delta^2 + (2\delta'Y - 2\delta'V\delta)\, n^{-\frac{1}{2}}$$

$$+ (3\delta'V^2\delta + Y'Y - 4\delta'VY)n^{-1} + R_{4n}$$

where R_{4n} is defined the same way as R_{1n}. Hence, for sufficiently large n and Y, V in J_n we expand $G_{2N}^{-\frac{1}{2}}$ and

$$G_{2N}^{-\frac{1}{2}} = \Delta^{-1} - \Delta^{-3}\, n^{-\frac{1}{2}}(\delta'Y - \delta'V\delta)$$

$$- n^{-1}[(2\Delta^3)^{-1}(3\delta'V^2\delta + Y'Y - 4\delta'VY)$$

$$- 3(2\Delta^5)^{-1}(\delta'Y - \delta'V\delta)^2] + R_{5n}$$

where R_{5n} is defined the same way as R_{1n}. Hence, for sufficiently large n and Y, Z, V in J_n, we have

$$(u\, a_N^{\frac{1}{2}} + G_{1N})\, G_{2N}^{-\frac{1}{2}} = u + n^{-\frac{1}{2}}\, \ell(Z,V) + n^{-1}\, Q(Y,Z,V) + R_{6n}$$

where R_{6n} is defined the same way as R_{3n}, and

$$\ell(Z,V) = (u/2\Delta^2)\delta'V\delta + \delta'Z/\Delta \tag{4.20}$$

$$Q(Y,Z,V) = \Delta^{-1}(Y'Z - \delta'VZ) \tag{4.21}$$

$$+ \Delta^{-2}\, u(\delta'VY - \delta'V^2\delta) - \Delta^{-3}(\delta'Y\delta'Z - \delta'Z\delta'V\delta)$$

$$+ u\,\Delta^{-4}[(7/8)(\delta'V\delta)^2 - \delta'Y\delta'V\delta]\ .$$

Hence, for sufficiently large n and Y, Z, V in J_n, we have

$$\Phi[(u\, a_N^{\frac{1}{2}} + G_{1N})\, G_{2N}^{-\frac{1}{2}}] = \Phi(u) + n^{-\frac{1}{2}}\, \varphi(u)\, \ell(Z,V) \tag{4.22}$$

$$+ n^{-1}\, \varphi(u)[Q(Y,Z,V) - (u/2)\ell^2(Z,V)] + R_{7n}$$

where R_{7n} is defined the same way as R_{3n}. We shall denote the three summands of R_{7n} as $R_{7n}(1)$, $R_{7n}(2)$ and $R_{7n}(3)$, respectively.

Let $g(Y,Z,V)$ be a function of Y,Z and V having finite expectation for g^2. Then

$$\begin{aligned}|E\, g(Y,Z,V) - E\, g(Y,Z,V)\, \chi(J_n)| &= |E\, g(Y,Z,V)\, \chi(J_n^c)| \\ &\le [E\, g^2(Y,Z,V)]^{\frac{1}{2}}\, [E\, \chi(J_n^c)]^{\frac{1}{2}} \\ &= o(n^{-1}) .\end{aligned}$$

Hence

$$\begin{aligned}&E\, \Phi[(u\, a_N^{\frac{1}{2}} + G_{1N})\, G_{2N}^{-\frac{1}{2}}]\, \chi(J_n) \\ (4.23)\quad &= \Phi(u) + n^{-\frac{1}{2}}\, \varphi(u)\, E[\ell(Z,V)\, \chi(J_n)] \\ &\quad + n^{-1}\, \varphi(u)\, E[Q(Y,Z,V) - (u/2)\ell^2(Z,V)] \\ &\quad + E[R_{7n}(1) + R_{7n}(2) + R_{7n}(3)]\, \chi(J_n) + o(n^{-2}) .\end{aligned}$$

But, $R_{7n}(1)$ has terms with zero expectation over J_n (since J_n is by definition symmetric in Y and Z) whenever Y, Z has odd power. Moreover, for any m

$$E\, V_{ij}^{2m-1} = n^{m-1/2}\, O(n^{-m}) = O(n^{-1/2}) .$$

Hence

$$(4.24)\qquad |E\, R_{7n}(1)\, \chi(J_n)| \le n^{-3/2}\, O(n^{-1/2}) = O(n^{-2}) .$$

Since fourth-order absolute moments of Y,Z,V exist and are bounded, we have

(4.25) $\quad |E\ R_{7n}(2)\ \chi(J_n)| \le E|R_{7n}(2)| = O(n^{-2})$.

Now

(4.26) $\quad |E\ R_{7n}(3)\ \chi(J_n)| = O(n^{-5/2}\ \log^5 n) = o(n^{-2})$

and

$$E\ \ell(Z,V)\ \chi(J_n) = (u/2\Delta^2)E\ \delta'V\delta\ \chi(J_n) + \Delta^{-1}\ E\ \delta'Z\ \chi(J_n) .$$

But $E\delta'Z\ \chi(J_n) = \delta'E(Z\ \chi(J_n)) = 0$ since $EZ = 0$ and J_n is symmetric in Z_i and the density of Z_i is symmetric about the origin, for $i = 1,\ldots,p$.

$$|E\ \delta'V\delta\ \chi(J_n)| = |\delta'[EV\ \chi(J_n)]\delta|$$

$$= |\delta'EV\ \chi(J_n^c)\delta|, \quad \text{since} \quad EV = 0 .$$

But

$$|E\ V_{ij}\chi(J_n^c)|$$

$$\le E|V_{ij}|\ \chi(J_n^c) = \text{Constant.}\ \ E|V_{ij}|\chi(|V_{ij}| > 2\log n)$$

$$\le \text{Constant.}\ [E|V_{ij}|^2]^{\frac{1}{2}}\ [P\{|V_{ij}| > 2\log n]^{\frac{1}{2}}$$

$$= O(n^{-k/2})$$

where k is given by (A.3). Since in (A.3) we could have taken any finite value of k, it follows that

(4.27) $\quad E\ \ell(Z,V)\ \chi(J_n) = E\ \ell(Z,V) + O(n^{-2})$.

Combining (4.24) through (4.27) we get the following theorem from (4.23):

Theorem 4.1.1.

When H_1 obtains,

$$P[W_a^* \le u] = \Phi(u) + n^{-1}\,\varphi(u)\,E[Q(Y,Z,V) - (u/2)\ell^2(Z,V)] + O(n^{-2}).$$

Replacing u by $(-v)$ and noticing that

$$P_1[(W-a_N)(4a_N)^{-\frac{1}{2}} \le u] = P_2[(W+a_N)(4a_N)^{-\frac{1}{2}} \ge -u] \tag{4.28}$$

where P_i is the probability under H_i, $i = 1,2$, we get an immediate corollary.

Corollary 4.1.1.

When H_2 obtains,

$$P[(W+a_N)(4a_N)^{-\frac{1}{2}} \le v]$$

$$= \Phi(v) - n^{-1}\,\varphi(v)E[Q(Y,Z,V) - (v/2)\ell^2(Z,V)] + O(n^{-2}).$$

Now we shall consider the three special cases (a), (b), and (c) as was introduced at the introduction of this section.

Case (a): Suppose we define $B \equiv I_p$ instead of $B = (B_1 + B_2)/2$ as given in (4.6), then V, as defined by (4.17) becomes identical to zero matrix. Thus $|V_{ij}| < 2\log n$ trivially holds for all i and j and hence lemma 4.1.1. is still applicable. The arguments used to get the expansion (4.22) are valid in this case with

$\ell(Z,V) \equiv \ell(Z,0)$, $Q(Y,Z,V) \equiv Q(Y,Z,0)$ and $R_{7n}(Y,Z,V) = R_{7n}(Y,Z,0)$.

Now

$$(4.29)\qquad E\,Q(Y,Z,0) = \Delta^{-1}\,EY'Z - \Delta^{-3}\,\delta'(EYZ')\delta$$

$$= \Delta^{-1}(n/N)\sum_{i=1}^{p}(1-\rho_i) - \Delta^{-3}\delta'(I_p - D)\delta(n/N)$$

$$= \Delta^{-1}(n/N)\mathrm{tr}(I_p - \Sigma^{-1}\Sigma_{12})$$

$$- \Delta^{-3}(n/N)[\Delta^2 - (\mu_1 - \mu_2)'\,\Sigma^{-1}\Sigma_{12}\Sigma^{-1}(\mu_1 - \mu_2)]\ .$$

$$(4.30)\qquad E\,\ell^2(Z,0) = \Delta^{-2}\,E(\delta'Z)^2 = (n/N)\ .$$

For Γ given by (4.2) with known Σ and known or unknown Σ_{12}, both ELR rule and PLR rule involve the statistic T_0 given by (3.2); we get the following theorem from (4.29), (4.30) and theorem 4.1.1. , for

$$a_{ON} = ((\bar{X}_1 - \bar{X}_2\ ;\ \Sigma\ ))$$

and replacing (n/N) by 2:

<u>Theorem 4.1.2.</u>

When H_1 obtains,

$$P[(T_0 - a_{ON}/2)\ a_{ON}^{-\frac{1}{2}} \le u]$$

$$= \Phi(u) + n^{-1}\,\varphi(u)[2\Delta^{-1}\,\mathrm{tr}(I_p - \Sigma^{-1}\Sigma_{12}) - 2\Delta^{-1} - u$$

$$+ 2\Delta^{-3}(\mu_1 - \mu_2)'\Sigma^{-1}\Sigma_{12}\Sigma^{-1}(\mu_1 - \mu_2)] + O(n^{-2}).$$

<u>Corollary 4.1.2.</u>

When H_2 obtains,

$$P[(T_0 + a_{ON}/2)\, a_{ON}^{-\frac{1}{2}} \le v]$$

$$= \Phi(v) - n^{-1}\,\varphi(v)[2\Delta^{-1}\,\mathrm{tr}(I_p - \Sigma^{-1}\Sigma_{12}) - 2\Delta^{-1} - v$$

$$+ 2\Delta^{-3}(\mu_1 - \mu_2)'\,\Sigma^{-1}\Sigma_{12}\Sigma^{-1}(\mu_1 - \mu_2)] + O(n^{-2}).$$

The following theorem is a special case of theorem 4.1.2 when $\Sigma_{12} = \rho\,\Sigma$ i.e. $D = \rho I_p$.

<u>Theorem 4.1.3.</u>

When H_1 obtains and $\Sigma_{12} = \rho\Sigma$,

$$P[(T_0 - a_{ON}/2)a_{ON}^{-\frac{1}{2}} \le u]$$

$$= \Phi(u) + n^{-1}\,\varphi(u)[2\Delta^{-1}(1-\rho)(p-1) - u] + O(n^{-2}) .$$

<u>Corollary 4.1.3.</u>

When H_2 obtains and $\Sigma_{12} = \rho\Sigma$,

$$P[(T_0 + a_{ON}/2)a_{ON}^{-\frac{1}{2}} \le v]$$

$$= \Phi(v) - n^{-1}\,\varphi(v)[2\Delta^{-1}(1-\rho)\,(p-1) - v] + O(n^{-2})$$

Before considering the remaining two cases (b) and (c), let us notice that

$$(4.31) \qquad E[Q(Y,Z,V) - (u/2)\,\ell^2(Z,V)]$$

$$= \Delta^{-1}(n/N)\,\mathrm{tr}(I_p - D) - \Delta^{-3}(n/N)\,\delta'(I_p - D)\delta$$

$$- (n/N)(u/2) - u\,\Delta^{-2}\,E(\delta' V^2 \delta)$$

$$+ [(7/8)\,u\,\Delta^{-4} - (u^3/8\Delta^4)]\,E\,(\delta' V \delta)^2$$

$$= \Delta^{-1}(n/N)\,\mathrm{tr}(I_p - \Sigma^{-1}\Sigma_{12}) - \Delta^{-3}(n/N)(\mu_1-\mu_2)'(\Sigma^{-1} - \Sigma^{-1}\Sigma_{12}\Sigma^{-1})(\mu_1-\mu_2)$$

$$- (u/2)(n/N) - u\,\Delta^{-2}\,E\,(\delta' V^2 \delta)$$

$$+ [(7/8)\,u\,\Delta^{-4} - (u^3/8\Delta^4)]\,E\,(\delta' V \delta)^2 .$$

Straightforward calculation shows that, if

$$M \sim W_p[\Lambda,\ m]$$

where Λ is a diagonal matrix with λ_i as the $i^{\underline{th}}$ diagonal entry, then

$$\text{(4.32)} \qquad E\,M = m\,\Lambda;\ E\,M^2 = \Lambda^*$$

where Λ^* is a diagonal matrix with $i^{\underline{th}}$ diagonal entry λ_i^* as given by

$$\text{(4.33)} \qquad \lambda_i^* = 2m\,\lambda_i^2 + m^2\,\lambda_i^2 + \sum_{\substack{j=1 \\ \neq i}}^{p} m\,\lambda_i\lambda_j .$$

<u>Case (b):</u> In this case Σ is unknown and $\Sigma_{12} = \rho\Sigma$ with known ρ and

$$V = n^{\frac{1}{2}}[(B_1 + B_2)/2 - I_p]$$

where B_1 and B_2 are given by (4.8). Hence it follows from (3.7) and (3.8) that

$$V = n^{-\frac{1}{2}}[M_1 + M_2 - nI_p]$$

where M_1 and M_2 are independent $W_p[I_p, n/2]$. From (4.32) and (4.33) we have, for $i = 1,2$

$$(4.34) \qquad E\, M_i = (n/2)I_p$$

$$E\, M_i^2 = (n + n^2/4 + (p-1)n/2)I_p \ .$$

Hence, from (4.34) and the fact that $(\delta'\delta)^{-1}\, \delta' M_i \delta$ has a chi-square distribution with $(n/2)$ degrees of freedom, we get

$$(4.35) \qquad E(\delta' V \delta)^2 = n^{-1}[2n + n^2/2 + n^2 + n^2/2 - 2n^2](\delta'\delta)^2$$

$$= 2\Delta^4$$

and

$$(4.36) \qquad E(\delta' V^2 \delta) = n^{-1}[2n + n^2/2 + (p-1)\, n + n^2 + n^2/2 - 2n^2]\delta'\delta$$

$$= (p + 1)\Delta^2 \ .$$

We can rewrite (4.31) using (4.35) and (4.36) as

$$E[Q(Y,Z,V) - (u/2)\ell^2(Z,V)]$$

$$= \Delta^{-1}(n/N)(p-1)(1-\rho) - u(p-3/4)$$

$$- u^3/4 - (u/2)(n/N) \ .$$

Replacing (n/N) by 2, it follows from Theorem 4.1.1 and for

$$a_{1N} = ((\overline{X}_1 - \overline{X}_2;\ (B_1 + B_2)/2))$$

that

Theorem 4.1.4. When H_1 obtains and $\Sigma_{12} = \rho\Sigma$,

$$P[(T_1 - a_{1N}/2)\, a_{1N}^{-\frac{1}{2}} \le u]$$

$$= \Phi(u) + n^{-1}\, \varphi(u)[2\, \Delta^{-1}(1-\rho)(p-1) - u(p + 1/4) - u^3/4]$$

$$+ O(n^{-2}) .$$

Corollary 4.1.4. When H_2 obtains and $\Sigma_{12} = \rho\Sigma$,

$$P[(T_1 + a_{1N}/2)\, a_{1N}^{-\frac{1}{2}} \le v]$$

$$= \Phi(v) - n^{-1}\, \varphi(v)[2\Delta^{-1}(1-\rho)(p-1) - v(p + 1/4) - v^3/4]$$

$$+ O(n^{-2}) .$$

Case (c): When both Σ and Σ_{12} are unknown, we have

$$V = n^{\frac{1}{2}}\, [(B_1 + B_2)/2 - I_p]$$

where B_1 and B_2 are given by (4.9). It follows from (2.7) and (3.6) that

$$V = (4n)^{-\frac{1}{2}}\, [M_3 + M_4 - 2nI_p]$$

where M_3 and M_4 are independently distributed and

$$M_3 \sim W_p[\Lambda_1,\ n/2]$$

$$M_4 \sim W_p[\Lambda_2,\ n/2]$$

where

$$\Lambda_1 = 2(I_p - D)$$
$$\Lambda_2 = 2(I_p + D) .$$

Now, it follows from (4.32) and (4.33) that

$$(4.37) \qquad E\ M_3 = (n/2)\Lambda_1 \ ;\ E\ M_3^2 = \Lambda_1^*$$

where Λ_1^* is a diagonal matrix with $i^{\underline{th}}$ diagonal entry λ_{1i}^* is given by

$$(4.38) \qquad \lambda_{1i}^* = 4n(1-\rho_i)^2 + n^2(1-\rho_i)^2 + \sum_{\substack{j=1\\ \neq i}}^{p} 2n(1-\rho_i)(1-\rho_j)$$

$$(4.39) \qquad E\ M_4 = (n/2)\Lambda_2 \ ;\ E\ M_4^2 = \Lambda_2^*$$

where Λ_2^* is a diagonal matrix with $i^{\underline{th}}$ diagonal entry λ_{2i}^* given by

$$(4.40) \qquad \lambda_{2i}^* = 4n(1+\rho_i)^2 + n^2(1-\rho_i)^2 + \sum_{\substack{j=1\\ \neq i}}^{p} 2n(1+\rho_i)(1+\rho_j) \ .$$

From (4.37) through (4.40) and the fact that $(\delta'\Lambda_1\delta)^{-1}\ \delta' M_3\delta$ and $(\delta'\Lambda_2\delta)^{-1}\delta' M_4\delta$ are independent chi-squares with (n/2) degrees of freedom, we get

$$E\ \delta' V^2\delta = (4n)^{-1}\ \delta'\Lambda_3\delta$$

where Λ_3 is a diagonal matrix with $i^{\underline{th}}$ diagonal entry λ_{3i} given by

$$\lambda_{3i} = 4n(1-\rho_i)^2 + n^2(1-\rho_i)^2 + \sum_{\substack{j=1\\ \neq i}}^{p} 2n(1-\rho_i)(1-\rho_j)$$

$$+ 4n(1+\rho_i)^2 + n^2(1+\rho_i)^2 + \sum_{\substack{j=1\\ \neq j}}^{p} 2n(1+\rho_i)(1+\rho_j)$$

$$+ 4n^2 + 2n^2(1-\rho_i)(1+\rho_i) - 4n^2(1-\rho_i) - 4n^2(1+\rho_i)$$

$$= 8n(1+\rho_i^2) + 4n \sum_{\substack{j=1 \\ \neq i}}^{p} (1+\rho_i\rho_j)$$

$$= 4n[(p+1) + \rho_i^2 + \rho_i(\sum_{j=1}^{p} \rho_i)]$$

$$= 4n[(p+1) + \rho_i^2 + \rho_i \text{ tr } D] .$$

Hence

$$(4.41) \qquad E(\delta'V^2\delta) = (p+1)\Delta^2 + \delta'D^2\delta + \delta'D\delta \text{ tr } D$$

$$= (p+1)\Delta^2 + (\mu_1-\mu_2)' \Sigma^{-1} \Sigma_{12} \Sigma^{-1} \Sigma_{12} \Sigma^{-1}(\mu_1-\mu_2)$$

$$+ (\mu_1-\mu_2)' \Sigma^{-1} \Sigma_{12} \Sigma^{-1}(\mu_1-\mu_2) \text{ tr } \Sigma^{-1} \Sigma_{12}$$

$$(4.42) \qquad E(\delta'V\delta)^2 = (4n)^{-1}[(\delta'\Lambda_1\delta)^2(n+n^2/4)+(\delta'\Lambda_2\delta)^2(n+n^2/4)$$

$$+ 4n^2 \Delta^4 + (n^2/2)\delta'\Lambda_1\delta\delta'\Lambda_2\delta$$

$$- 2n^2 \delta'\Lambda_1\delta \Delta^2 - 2n^2 \Delta^2 \delta'\Lambda_2\delta]$$

$$= [(\delta'\Lambda_1\delta)^2 + (\delta'\Lambda_2\delta)^2]/4$$

$$= (\Delta^2 - \delta'D\delta)^2 + (\Delta^2 + \delta'D\delta)^2$$

$$= 2[\Delta^4 + \{(\mu_1-\mu_2)' \Sigma^{-1} \Sigma_{12} \Sigma^{-1}(\mu_1-\mu_2)\}^2] .$$

For $a_{2N} = ((\overline{X}_1 - \overline{X}_2; (2n)^{-1}(S_{11}^* + S_{22}^*)))$ and replacing (n/N)

by 2, we get from (4.31), (4.41), (4.42) and theorem 4.1.1 the following theorem:

<u>Theorem 4.1.5.</u> When H_1 obtains,

$$
\begin{aligned}
&P[(T_2 - a_{2N}/2)\, a_{2N}^{-\frac{1}{2}} \le u] \\
&\quad = \Phi(u) + n^{-1}\varphi(u)[2\Delta^{-1}\operatorname{tr}(I_p - \Sigma^{-1}\Sigma_{12}) - u \\
&\qquad - 2\Delta^{-3}(\mu_1-\mu_2)'(\Sigma^{-1} - \Sigma^{-1}\Sigma_{12}\Sigma^{-1}\Sigma_{12}\Sigma^{-1})(\mu_1-\mu_2) \\
&\qquad - u(p+1) - u\,\Delta^{-2}(\mu_1-\mu_2)'\Sigma^{-1}\Sigma_{12}\Sigma^{-1}\Sigma_{12}\Sigma^{-1}(\mu_1-\mu_2) \\
&\qquad - u\,\Delta^{-2}(\mu_1-\mu_2)'\Sigma^{-1}\Sigma_{12}\Sigma^{-1}(\mu_1-\mu_2)\operatorname{tr}\Sigma^{-1}\Sigma_{12} \\
&\qquad + 7u/4 - u^3/4 \\
&\qquad + \Delta^{-4}(7u/4 - u^3/4)\{(\mu_1-\mu_2)'\Sigma^{-1}\Sigma_{12}\Sigma^{-1}(\mu_1-\mu_2)\}^2] \\
&\qquad + O(n^{-2}) .
\end{aligned}
$$

<u>Corollary 4.1.5.</u> When H_2 obtains,

$$
\begin{aligned}
&P[(T_2 + a_{2N}/2)a_{2N}^{-\frac{1}{2}} \le v] \\
&= \Phi(v) - n^{-1}\varphi(v)[2\Delta^{-1}\operatorname{tr}(I_p - \Sigma^{-1}\Sigma_{12}) - v \\
&\qquad - 2\Delta^{-3}(\mu_1-\mu_2)'(\Sigma^{-1} - \Sigma^{-1}\Sigma_{12}\Sigma^{-1}\Sigma_{12}\Sigma^{-1})(\mu_1-\mu_2) \\
&\qquad - v(p+1) - v\,\Delta^{-2}(\mu_1-\mu_2)'\Sigma^{-1}\Sigma_{12}\Sigma^{-1}\Sigma_{12}\Sigma^{-1}(\mu_1-\mu_2) \\
&\qquad - v\,\Delta^{-2}(\mu_1-\mu_2)'\Sigma^{-1}\Sigma_{12}\Sigma^{-1}(\mu_1-\mu_2)\operatorname{tr}\Sigma^{-1}\Sigma_{12}
\end{aligned}
$$

$$+ (7v/4 - v^3/4)(1 + \Delta^{-4}\{(\mu_1-\mu_2)' \Sigma^{-1} \Sigma_{12} \Sigma^{-1}(\mu_1-\mu_2)\}^2)]$$

$$+ O(n^{-2})$$

We get the following corollary from Theorem 4.1.5.

<u>Corollary 4.1.5.</u> When H_1 obtains and $\Sigma_{12} = \rho\Sigma$,

$$P[(T_2 - a_{2N}/2)\, a_{2N}^{-\frac{1}{2}} \le u]$$

$$= \Phi(u) + n^{-1} \varphi(u)[2\Delta^{-1}(1-\rho)(p-1) - u(1+\rho^2)(p-3/4)$$

$$- u - (u^3/4)(1 + \rho^2)] + O(n^{-2}) .$$

<u>Corollary 4.1.7.</u> When H_2 obtains and $\Sigma_{12} = \rho\Sigma$,

$$P[(T_2 + a_{2N}/2)\, a_{2N}^{-\frac{1}{2}} \le v]$$

$$= \Phi(v) - n^{-1} \varphi(v)[2\Delta^{-1}(1-\rho)(p-1) - v(1+\rho^2)(p-3/4)$$

$$- v - (v^3/4)(1+\rho^2)] + O(n^{-2}) .$$

4.2. <u>Asymptotic expansion of the distribution function of W_α^*</u>.

We shall now consider W_α^* given by (4.4) and assume (4.10) since W_α^* is invariant under the transformations given by (4.11). The distribution function of W_α^* is given by, under H_1,

$$P[W_\alpha^* \le u] = P[U_0 \le (\omega + \tfrac{1}{2}G_{3N})G_{2N}^{-\frac{1}{2}}]$$

where $\omega = u\,\alpha^{\frac{1}{2}} + \alpha/2 = u\,\Delta + \Delta^2/2$, G_{2N} is given by (4.13) and

$$G_{3N} = (\bar{X}_1 + \bar{X}_2)' \; B^{-1}(\bar{X}_1 - \bar{X}_2) \; .$$

Hence

$$P[W_\alpha^* \le u] = E \; \Phi[(\omega + \tfrac{1}{2} \, G_{3N}) \; G_{2N}^{-\frac{1}{2}}]$$

where expectation is taken over the joint density of $\bar{X}_1$, $\bar{X}_2$ and B.

As we have done in Section 4.2, let $n = 2(N-1)$ and define Y, U and V by

$$(4.43) \qquad (\bar{X}_1 + \bar{X}_2) = U \, n^{-\frac{1}{2}} - \delta$$

and Y and V are given by (4.16) and (4.17) respectively.

Then

$$\begin{pmatrix} Y \\ U \end{pmatrix} \sim N_{2p}\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \begin{pmatrix} 2(n/N)(I_p - D) & 0 \\ 0 & 2(n/N)(I_p + D) \end{pmatrix}\right].$$

Hence we also have Y and U independent. We shall now redefine J_n as J_n^* given by

$$(4.44) \qquad J_n^* = \{|Y_i| < 4(\log n)^{\frac{1}{2}} \; , \quad |U_i| < 4(\log n)^{\frac{1}{2}},$$

$$|V_{ij}| < 2 \log n \; ; \; i,j = 1,\ldots,p\} \; .$$

It follows from the proof of Lemma 4.1.1 that

$$P(J_n^*) = 1 - o(n^{-2}).$$

Now, following (4.19) through (4.27) we get

$$P[W^* \le u] = \Phi(u) + \varphi(u)n^{-\frac{1}{2}} E\ \ell^*(Y,U,V)$$

$$+ \varphi(u)n^{-1} E[Q^*(Y,U,V) - (u/2)\ \ell^{*2}(Y,U,V)] + O(n^{-2})$$

where

$$\ell^*(Y,U,V) = (2\Delta)^{-1}(U'\delta+\delta'V\delta-\delta'Y) - u\ \Delta^{-2}(\delta'Y-\delta'V\delta) \tag{4.45}$$

and

$$\begin{aligned} Q^*(Y,U,V) = {} & (2\Delta)^{-1}\ [U'Y+\delta'VY-\delta'V^2\delta-U'V\delta] \\ & - (2\Delta^3)^{-1}\ [(\delta'Y-\delta'V\delta)(U'\delta+\delta'V\delta-\delta'Y)] \\ & - u\ [(2\Delta^2)^{-1}(3\delta'V^2\delta+Y'Y-4\delta'VY) \\ & - 3(2\Delta^4)^{-1}(\delta'Y-\delta'V\delta)^2]\ . \end{aligned} \tag{4.46}$$

Now $E\ \ell^*(Y,U,V) = 0$

$$\begin{aligned} E\ Q^*(Y,U,V) = {} & - (2\Delta)^{-1}\ E\delta'V^2\delta \\ & + (2\Delta^3)^{-1}\ E[(\delta'Y)^2 + (\delta'V\delta)^2] \\ & - u\ E[(2\Delta^2)^{-1}(3\delta'V^2\delta + Y'Y) \\ & - 3(2\Delta^4)^{-1}\ \{(\delta'Y)^2 + (\delta'V\delta)^2\}] \\ = {} & E\ \delta'V^2\delta[- (2\Delta)^{-1} - 3u(2\Delta^2)^{-1}] \\ & + E(\delta'V\delta)^2\ [(2\ \Delta^3)^{-1} + 3u(2\Delta^4)^{-1} \\ & + (2\Delta^3)^{-1}\ 2(n/N)\ \delta'(I_p - D)\delta \\ & - u(2\Delta^2)^{-1}\ 2(n/N)\ tr\ (I_p - D) \\ & + 3u(2\Delta^4)^{-1}\ 2(n/N)\ \delta'(I_p - D)\delta \end{aligned} \tag{4.47}$$

$$= E(\delta' V\delta)^2 \, [(2\Delta^3)^{-1} + 3u(2\Delta^4)^{-1}]$$

$$- E(\delta' V^2 \delta)[(2\Delta)^{-1} + 3u(2\Delta^2)^{-1}]$$

$$+ (2\Delta^3)^{-1} \, 2(n/N)(\mu_1-\mu_2)' \, (\Sigma^{-1} - \Sigma^{-1} \Sigma_{12} \Sigma^{-1})(\mu_1-\mu_2)$$

$$- u(2\Delta^2)^{-1} \, 2(n/N) \text{ tr } (I_p - \Sigma^{-1} \Sigma_{12})$$

$$+ 3u(2\Delta^4)^{-1} \, 2(n/N)(\mu_1-\mu_2)' \, (\Sigma^{-1} - \Sigma^{-1} \Sigma_{12} \Sigma^{-1})(\mu_1-\mu_2)$$

(4.48)
$$
\begin{aligned}
E \, \ell^{*2}(Y,U,V) &= (2\Delta)^{-2} \, E[\delta' UU'\delta + \delta' YY'\delta + (\delta' V\delta)^2] \\
&\quad + u^2 \, \Delta^{-4} \, E[\delta' YY'\delta + (\delta' V\delta)^2] \\
&\quad + u \, \Delta^{-3} \, E[\delta' YY'\delta + (\delta' V\delta)^2] \\
&= (2\Delta)^{-2} \, 2(n/N) \, [\delta' \, (I_p + D)\delta + \delta'(I_p - D)\delta] \\
&\quad + (u^2 \, \Delta^{-4} + u \, \Delta^{-3}) \, 2(n/N)\delta' \, (I_p - D)\delta \\
&\quad + E(\delta' V\delta)^2 \, [(2\Delta)^{-2} + u^2 \, \Delta^{-4} + u \, \Delta^{-3}] \\
&= (n/N) + (u^2 \, \Delta^{-4} + u \, \Delta^{-3}) \, 2(n/N) \\
&\cdot (\mu_1 - \mu_2)'(\Sigma^{-1} - \Sigma^{-1} \Sigma_{12} \Sigma^{-1})(\mu_1-\mu_2) \\
&\quad + [(2\Delta)^{-2} + u^2\Delta^{-4} + u \, \Delta^{-3}] \, E \, (\delta' V\delta)^2 \, .
\end{aligned}
$$

Hence, we can obtain similar theorems as was given in Section 4.2.

Theorem 4.2.1. When H_1 obtains,

$$P[W_\alpha^* \le u] = \Phi(u) + \varphi(u) \, n^{-1} \, E[Q^*(Y,U,V) - (u/2)\ell^{*2}(Y,U,V)]$$
$$+ O(n^{-2}) \, .$$

Now we consider the three cases (a), (b) and (c).

Case (a): In the case when Σ is known, we take $B = I_p$ and V is a zero matrix. Hence, arguing as before, we get the following theorem, for T_0 given by (3.2) and replacing (n/N) by 2:

Theorem 4.2.2. When H_1 obtains,

$$
\begin{aligned}
&P[(T_0 - \alpha/2)\,\alpha^{-\frac{1}{2}} \le u] \\
&\quad = \Phi(u) + n^{-1}\,\varphi(u)[2\Delta^{-3}\,(\mu_1-\mu_2)'(\Sigma^{-1} - \Sigma^{-1}\Sigma_{12}\Sigma^{-1})(\mu_1-\mu_2) \\
&\qquad - u - 2u\,\Delta^{-2}\,\mathrm{tr}(I_p - \Sigma^{-1}\Sigma_{12}) \\
&\qquad - 2(u^3\Delta^{-4} + u^2\Delta^{-3})(\mu_1-\mu_2)'(\Sigma^{-1} - \Sigma^{-1}\Sigma_{12}\Sigma^{-1})(\mu_1-\mu_2) \\
&\qquad + 6u\,\Delta^{-4}(\mu_1-\mu_2)'(\Sigma^{-1} - \Sigma^{-1}\Sigma_{12}\Sigma^{-1})(\mu_1-\mu_2)] + O(n^{-2}) .
\end{aligned}
$$

Corollary 4.2.1. When H_1 obtains and $\Sigma_{12} = \rho\Sigma$,

$$
\begin{aligned}
&P[(T_0 - \alpha/2)\alpha^{-\frac{1}{2}} \le u] \\
&\quad = \Phi(u) + n^{-1}\,\varphi(u)[2(1-\rho)\Delta^{-1} - u - 2(p-3)(1-\rho)\,u\Delta^{-2} \\
&\qquad - 2(1-\rho)(u^3\Delta^{-2} + u^2\Delta^{-1})] + O(n^{-2}) .
\end{aligned}
$$

Case (b): In this case Σ is unknown and $\Sigma_{12} = \rho\Sigma$ with known ρ. Hence from (4.35), (4.36) and theorem 4.2.1 we get the following theorem, replacing (n/N) by 2 and for T_1 given by (3.4):

Theorem 4.2.3. When H_1 obtains and $\Sigma_{12} = \rho\Sigma$,

$$P[(T_1 - \alpha/2)\,\alpha^{-\frac{1}{2}} \le u]$$

$$= \Phi(u) + n^{-1}\,\varphi(u)[2(1-\rho)\,\Delta^{-1} - u(3p-1)/2 - \Delta(p-1)/2$$

$$- 2(p-3)(1-\rho)\,u\,\Delta^{-2} - 2(1-\rho)(u^3\,\Delta^{-2} + u^2\,\Delta^{-1})$$

$$- u(\Delta^2/4 + u^2 + u\Delta)] + O(n^{-2})\ .$$

<u>Case (c)</u>: When both Σ and Σ_{12} are unknown we get the following theorem from (4.42), (4.43) and theorem 4.2.1, replacing (n/N) by 2 and for T_2 given by (3.23)

<u>Theorem 4.2.4.</u> When H_1 obtains ,

$$P[(T_2 - \alpha/2)\alpha^{-\frac{1}{2}} \le u]$$

$$= \Phi(u) + n^{-1}\,\varphi(u)\Big[2\Delta^{-3}(\mu_1-\mu_2)'(\Sigma^{-1} - \Sigma^{-1}\Sigma_{12}\Sigma^{-1})(\mu_1-\mu_2)$$

$$- u - 2u\,\Delta^{-2}\,\mathrm{tr}(I_p - \Sigma^{-1}\Sigma_{12})$$

$$- 2(u^3\Delta^{-4} + u^2\Delta^{-3})(\mu_1-\mu_2)'(\Sigma^{-1} - \Sigma^{-1}\Sigma_{12}\Sigma^{-1})(\mu_1-\mu_2)$$

$$+ 6u\,\Delta^{-4}(\mu_1-\mu_2)'(\Sigma^{-1} - \Sigma^{-1}\Sigma_{12}\Sigma^{-1})(\mu_1-\mu_2)$$

$$+ 2\{\Delta^4 + [(\mu_1-\mu_2)'\,\Sigma^{-1}\Sigma_{12}\Sigma^{-1}(\mu_1-\mu_2)]^2\}\{(2\Delta^3)^{-1} + 3u(2\Delta^4)^{-1}\}$$

$$- \{(2\Delta)^{-1} + 3u(2\Delta^2)^{-1}\}$$

$$\{(p+1)\Delta^2 + (\mu_1 - \mu_2)'\,\Sigma^{-1}\Sigma_{12}\Sigma^{-1}\Sigma_{12}\Sigma^{-1}\,(\mu_1 - \mu_2)$$

$$+ (\mu_1-\mu_2)'\,\Sigma^{-1}\Sigma_{12}\Sigma^{-1}(\mu_1-\mu_2)\,\mathrm{tr}\,\Sigma^{-1}\Sigma_{12}\}$$

$$- (u/2)\{(2\Delta)^{-2} + u^2\Delta^{-4} + u\Delta^{-3}\}$$

$$\cdot\,\{2\Delta^4 + 2[(\mu_1 - \mu_2)'\,\Sigma^{-1}\Sigma_{12}\Sigma^{-1}\,(\mu_1 - \mu_2)]^2\}\Big] + O(n^{-2})\ .$$

As a special case of theorem 4.2.4 we get the following:

Corollary 4.2.2. When H_1 obtains and $\Sigma_{12} = \rho\Sigma$,

$$P[(T_2 - \alpha/2)\,\alpha^{-\frac{1}{2}} \le u]$$

$$= \Phi(u) + n^{-1}\,\varphi(u)[2(1-\rho)\Delta^{-1} - u - 2(p-3)(1-\rho)u\,\Delta^{-1}$$

$$- 2(1-\rho)(u^3\Delta^{-2} + u^2\Delta^{-1}) + (2\Delta^4 + 2\Delta^4\rho^2)\{(2\Delta^3)^{-1} + 3u(2\Delta^4)^{-1}\}$$

$$- \{(2\Delta)^{-1} + 3u(2\Delta^2)^{-1}\}\,\Delta^2(p+1)(1+\rho^2)$$

$$- u\,\{(2\Delta)^{-2} + u^2\Delta^{-4} + u\,\Delta^{-3}\}(\Delta^4 + \Delta^4\rho^2)] + O(n^{-2}).$$

Similar theorems and corollaries can be written when H_2 obtains using the following relation:

$$(4.49) \qquad P_1[W_\alpha^* \le u] = P_2[(W + \alpha)(4\alpha)^{-\frac{1}{2}} \ge -u]$$

where P_i is the probability under H_i, $i = 1,2$.

4.3. Asymptotic expansion of probability of misclassification of some classification rules:

Let us consider the rule which classifies X_0 to π_1 if, and only if,

$$(4.50) \qquad W \ge 0$$

where W is given by (4.1). It follows from (4.49) and the definition of W_α^* that the probability of misclassification (PMC) of the rule given by (4.50), under both H_1 and H_2, is

$$(4.51) \qquad P[W < 0] = P[W_\alpha^* < -\Delta/2]\ .$$

Now we consider the three cases (a), (b), and (c) given before.

Case (a): In this case Σ is known and Σ_{12} may be known or unknown. We have seen in Section 2 that both the EML rule and the PML rule accept H_1 if, and only if

$$T_0 \geq 0$$

where T_0 is given by (3.2). Hence we get the following theorem from (4.51) and theorem 4.2.2.

Theorem 4.3.1. When Σ is known, the PMC of the EML rule as well as the PML rule is

$$P[T_0 < 0] = \Phi(-\Delta/2) + n^{-1}\varphi(\Delta/2)\ [\Delta/2 + \Delta^{-1}\operatorname{tr}(I_p - \Sigma^{-1}\Sigma_{12}) - (\mu_1-\mu_2)'(\Sigma^{-1} - \Sigma^{-1}\Sigma_{12}\Sigma^{-1})(\mu_1-\mu_2)\{\Delta^{-3} + (4\Delta)^{-1}\}] + O(n^{-2})\ .$$

Corollary 4.3.1. When Σ is known and $\Sigma_{12} = \rho\Sigma$, the PMC of the EML rule as well as the PML rule is

$$P[T_0 < 0] = \Phi(-\Delta/2) + n^{-1}\varphi(\Delta/2)[\Delta(1+\rho)/4 + \Delta^{-1}(p-1)(1-\rho)] + O(n^{-2}).$$

Case (b): In this case Σ is unknown and $\Sigma_{12} = \rho\Sigma$ with known ρ. We have seen in Section 2 that both the EML rule and the PML rule accept H_1 if, and only if

$$T_1 \geq 0$$

where T_1 is given by (3.4). Hence we get the following theorem from (4.51) and theorem 4.2.3.

Theorem 4.3.2. When Σ is unknown and $\Sigma_{12} = \rho\Sigma$ with know ρ, the PMC of the PML as well as the EML rule is

$$P[T_2 < 0] = \Phi(-\Delta/2) + n^{-1}\varphi(\Delta/2)[\Delta(p-1)/4 + \Delta(1+\rho)/4 + (p-1)(1-\rho)\Delta^{-1}] + O(n^{-2}) .$$

Case (c): In this case both Σ and Σ_{12} are unknown. We have seen in Section 2 that the PML rule accepts H_1 if, and only if,

$$T_2 \geq 0$$

where T_2 is given by (3.23). Hence we get the following theorem from (4.51) and theorem 4.2.4.

Theorem 4.3.3. When both Σ and Σ_{12} are unknown. the PMC of the PML rule is

$$\begin{aligned} P[T_2 < 0] = \Phi(-\Delta/2) + n^{-1}\varphi(\Delta/2)[\Delta/2 + \Delta^{-1}\operatorname{tr}(I_p - \Sigma^{-1}\Sigma_{12}) \\ - (\mu_1-\mu_2)'(\Sigma^{-1} - \Sigma^{-1}\Sigma_{12}\Sigma^{-1})(\mu_1-\mu_2)\{\Delta^{-3} + (4\Delta)^{-1}\} \\ - (2\Delta^3)^{-1}\{\Delta^4 + [(\mu_1-\mu_2)'\Sigma^{-1}\Sigma_{12}\Sigma^{-1}(\mu_1-\mu_2)]^2\} \\ + (4\Delta)^{-1}\{(p+1)\Delta^2 + (\mu_1-\mu_2)'\Sigma^{-1}\Sigma_{12}\Sigma^{-1}\Sigma_{12}\Sigma^{-1}(\mu_1-\mu_2) \\ + (\mu_1-\mu_2)'\Sigma^{-1}\Sigma_{12}\Sigma^{-1}(\mu_1-\mu_2)\operatorname{tr}(\Sigma^{-1}\Sigma_{12})\}] + O(n^{-2}) .\end{aligned}$$

We have also seen in Section 2 that when $\Sigma_{12} = \rho\Sigma$, the PML rule for $p = 1$ and the approximate PML rule for $p > 1$ are the same as the rule for general Σ_{12}. Hence we get the following theorem from (4.51) and corollary 4.2.2 (or theorem 4.3.3).

Theorem 4.3.4. When $\Sigma_{12} = \rho\Sigma$ and both Σ and ρ are unknown, the PMC of the approximate (exact when $p = 1$) PML rule is

$$P[T_2 < 0] = \Phi(-\Delta/2) + n^{-1}\varphi(\Delta/2)[\Delta(1+\rho)/4$$
$$+ \Delta(1+\rho^2)(p-1)/4 + (p-1)(1-\rho)\Delta^{-1}] + 0(n^{-2}) .$$

Remarks: (i). We have seen in Section 3 that the computation of the percentage points from the exact distribution of W is not practical; results of this section may be used to compute the percentage points (approximate to the order of n^{-2}). For an analysis in this line, when Σ_{12} is a zero matrix, one may refer to Anderson [4].

(ii). As Δ tends to infinity, the PMC (to the order of n^{-2}) in all the cases tend to $\Phi(-\Delta/2)$ in such a way that

$$\lim_{\Delta\to\infty} [\mathrm{PMC} - \Phi(-\Delta/2)]/\Phi(-\Delta/2) = 0 .$$

(iii). When $p = 1$, large negative values of ρ seem to give better approximation since the PMC of all the rules discussed above decrease as ρ tends to -1. However, in Section 5 we have established that the exact PMC decreases as ρ tends to -1 if only Δ is large and increases as ρ tends to -1 if Δ is small; bounds of ρ for different combinations of Δ and N have also been obtained.

(iv). When $p > 1$ and $\Sigma_{12} = \rho\Sigma$, from corollary 4.3.1 and theorem 4.3.2 we observe that large negative ρ gives a better approximation if and only if $\Delta^2 > 4(p-1)$. If $\Delta^2 < 4(p-1)$, the approximation is better when ρ is close to 1; and if $\Delta^2 = 4(p-1)$, the approximate PMC does not involve ρ. From theorem 4.3.4 it follows that the PMC to the order of n^{-2} is minimum (for fixed N and Δ) at $\rho = 2\Delta^{-2} - (2p - 2)^{-1}$. Hence, in this case $\rho = 0$ or the standard case gives best approximation if $\Delta^2 = 4(p-1)$. Otherwise $\rho = 2\Delta^{-2} - (2p-2)^{-1}$ gives a better approximation if

$$|2\Delta^{-2} - (2p - 2)^{-1}| \leq 1.$$

In case $2\Delta^{-2} - (2p - 2)^{-1} > 1$, $\rho = 1$ gives a better approximation and $\rho = -1$ otherwise.

5. STUDIES ON PROBABILITY OF CORRECT CLASSIFICATION.

5.0. Introduction.

In this section we shall study the probability of correct classification (PCC) of some rules as a function of the unknown parameters and two estimates of the PCC of the ML rule and inequalities concerning their expectations and PCC of the EML rule. In the Section 5.1. we take Γ as given by (3.1).

Consider the class of rules ψ_k which accepts H_1 if, and only if,

$$W \geq k \tag{5.1}$$

for some constant k; where W is given by (4.1). Let ψ_{ik} be the class of rules which accepts H_1 if, and only if,

$$T_i \geq k \tag{5.2}$$

for some constant k, $(i = 0,1,2)$; where T_o, T_1 and T_2 are given respectively by (3.2), (3.4) and (3.23). We have seen in section 4 that ψ_{ik}, $i = 0,1,2$ are special cases of ψ_k since T_i, $i = 0,1,2$ are special cases of W for $2B = B_1 + B_2$ given, respectively by (4.7), (4.8) and (4.9). We have also seen that W is invariant under the transofrmations (4.11). But, if we assume (3.1), we can take L in (4.11) such that

$$L \Sigma L' = I_p, \quad L(\mu_1 - \mu_2) = (\Delta, 0, \ldots, 0) = \delta' \tag{5.3}$$

and as a consequence the PCC of the rule ψ_k, and hence the rules

ψ_{ik}, $i = 0,1,2$, involve only two parameters, α and ρ, where $\alpha = \Delta^2$ is given by (3.13). In view of (4.11), we shall take $\Sigma = I_p$, $\mu_1 = 0$ and $\mu_2' = (\Delta, 0,\dots,0)$, under the condition (3.1). In the first part of this section we shall study the monotonicity of PCC of ψ_k as a function of Δ, for fixed ρ. In the second part we shall consider $p = 1$ (in which case (3.1) is trivially true) and study the PCC of ψ_0 as a function of ρ for fixed Δ. In each case we shall hold N fixed. Finally we shall consider the estimation problem of the PCC of the ML rule, for small as well as large N and inequalities concerning PCC and the estimates.

5.1. Monotonicity of the PCC of the rule ψ_k as a function of α for fixed ρ and N.

We shall use the following result (with different notation) of Das Gupta [11] to show that the PCC of the rule ψ_k increases monotonically with α for fixed ρ and N when H_1 obtains and $k \geq 0$.

Das Gupta's Result: Let U_1 and U_2 be two $p \times 1$ random vectors such that

$$\begin{pmatrix} U_1 \\ U_2 \end{pmatrix} \sim N_{2p}\left[\begin{pmatrix} \nu \\ a\nu \end{pmatrix}, I_{2p}\right].$$

Let V be a $p \times p$ random matrix, distributed independently of U_1 and U_2, such that, for every $p \times p$ orthogonal matrix L, V and LVL' have the same distribution and the distribution of V does not involve ν. Then, both

$$P[U_1' U_1 \leq c] \quad \text{and} \quad P[U_1' V^{-1} U_2 \leq c]$$

depend on ν only through $\nu'\nu$ and they are increasing monotonically as a function of $\nu'\nu$ when $a \leq 0$ and $c \leq 0$.

For the rule ψ_k, we identify U_1, U_2 V, ν, a and c as follows, under H_1,

$$U_1 \equiv (\overline{X}_1 - \overline{X}_2)c_1^{-\frac{1}{2}}$$

$$U_2 \equiv (\overline{X}_1 + \overline{X}_2 - 2X_0)c_2^{-\frac{1}{2}}$$

$$\nu \equiv \delta, \; \nu'\nu = \delta'\delta = \alpha$$

$$a = -1, \quad V \equiv B, \quad c = -k(c_1c_2)^{-\frac{1}{2}}$$

where c_1 and c_2 are given by (3.3). Now, the PCC of the rule ψ_k, when H_1 obtains, is

$$P_1[W \geq k] = P_1[U_1' B^{-1} U_2 \leq -k(c_1c_2)^{-\frac{1}{2}}] .$$

Hence we get the following theorem:

<u>Theorem 5.1.1.</u> For fixed ρ and N, when H_1 obtains, the PCC of the rule ψ_k is an increasing function of α if $k \geq 0$.

<u>Corollary 5.1.1.</u> When H_1 obtains and ρ and N are held fixed, the PCC's of each of the rules ψ_{0k}, ψ_{1k} and ψ_{2k} are increasing functions of α if $k \geq 0$.

In view of (4.53) we get the following corollary:

Corollary 5.1.2. When either of H_1 or H_2 obtains and ρ and N are held fixed, the PCC's of each of the rules ψ_{00}, ψ_{10} and ψ_{20} are increasing functions of α.

Remark: The rules ψ_{00}, ψ_{10} and ψ_{20} are, respectively, the EML (or PML) rule when Σ is known, the EML (or PML) rule when Σ is unknown but ρ is known and the approximate (exact when $p = 1$) PML rule when both Σ and ρ are unknown. For $p = 1$ and fixed N, table 1 gives the PCC of the rule ψ_0 as a function of Δ for three different values of ρ, namely $\rho = 0.9999$, $\rho = 0$ and $\rho = -0.9999$.

5.2. PCC as a function of ρ when $p = 1$ and N and Δ are fixed.

When $p = 1$, the rules ψ_0, ψ_{i0}, $i = 0,1,2$ reduce to the same rule which accepts H_1 if, and only if

$$(\bar{X}_2 - \bar{X}_1)(2X_0 - \bar{X}_1 - \bar{X}_2) \le 0 .$$

The PCC of the above rule under both H_1 and H_2 are equal and is given by

$$h(\rho;\, \Delta,N) \equiv 1 - \Phi(\Delta\, c_1^{-\frac{1}{2}}) - \Phi(\Delta\, c_2^{-\frac{1}{2}}) + 2\Phi(\Delta\, c_1^{-\frac{1}{2}})\Phi(\Delta\, c_2^{-\frac{1}{2}}) \tag{5.4}$$

where c_1 and c_2 are given by (3.3).

Differentiating h partially with respect to ρ we get

$$\begin{aligned} h'(\rho;\, \Delta,N) &\equiv \frac{\partial h(\rho;\, \Delta,N)}{\partial\rho} \\ &= [2\Phi(\Delta\, c_2^{-\frac{1}{2}}) - 1]\varphi(\Delta\, c_1^{-\frac{1}{2}})\Delta\, c_1^{-3/2}\, N^{-1} \\ &\quad - [2\Phi(\Delta\, c_1^{-\frac{1}{2}}) - 1]\varphi(\Delta\, c_2^{-\frac{1}{2}})\Delta\, c_2^{-3/2}\, N^{-1}. \end{aligned} \tag{5.5}$$

Hence $h'(\rho;\ \Delta,N) \geq 0$ is equivalent to

$$(5.6)\qquad \frac{2\Phi(\Delta\ c_2^{-\frac{1}{2}}) - 1}{\varphi(\Delta\ c_2^{-\frac{1}{2}})}\ c_2^{-3/2} \geq \frac{2\Phi(\Delta\ c_1^{-\frac{1}{2}}) - 1}{\varphi(\Delta\ c_1^{-\frac{1}{2}})}\ c_1^{-3/2} .$$

We shall exclude the case when $\Delta \equiv 0$, since in this case $h(\rho;\ \Delta,N) \equiv 1/2$.

Define

$$(5.7)\qquad g(c) = [2\Phi(c) - 1]/\varphi(c)c^3, \qquad c > 0$$

Let

$$(5.8)\qquad \begin{aligned} g_1(\rho;\ \Delta,N) &\equiv g(\Delta\ c_1^{-\frac{1}{2}}) \\ g_2(\rho;\ \Delta,N) &\equiv g(\Delta\ c_2^{-\frac{1}{2}}) . \end{aligned}$$

Then we can rewrite (5.6) as

$$(5.9)\qquad g_2(\rho;\ \Delta,N) \geq g_1(\rho;\ \Delta,N) .$$

Note that $g(c) \to \infty$ as $c \to 0$. Thus, as $\rho \to 1$

$$(5.10)\qquad \begin{aligned} g_1(\rho;\ \Delta,N) &\to \infty \\ g_2(\rho;\ \Delta,N) &\to g(\Delta(4 + 4/N)^{-\frac{1}{2}}). \end{aligned}$$

Moreover,

$$(5.11)\qquad \begin{aligned} g_1(-1;\ \Delta,N) &= g(n^{\frac{1}{2}}\ \Delta/2) \\ g_2(-1;\ \Delta,N) &= g(\Delta/2) . \end{aligned}$$

Numerical calculations of g indicate that g is decreasing on $(0,C^*]$ and increasing on $[C^*,\infty)$, where $C^* = 1.6151575$, correct to 7 decimal places. Numerical calculations of the values of g_1 and g_2 for different combinations of Δ and N and also the graphs of $h(\rho;\ \Delta,N)$ give evidence that the following proposition is true (see TABLE 2).

Proposition 5.2.1. For fixed Δ and N,

$$g_1(\rho;\ \Delta,N) - g_2(\rho,\Delta,N)$$

can be equal to zero at most once, for $|\rho| \leq 1$.

In this sequel, we shall argue and make conclusions, assuming that this proposition is indeed true. Since $g_1 - g_2$ is a continuous function of ρ $(-1 \leq \rho < 1)$, it follows from proposition 5.2.1. and (5.10) that

$$g_1(\rho;\ \Delta,N) - g_2(\rho;\ \Delta,N) \geq 0$$

for any ρ if

$$g_1(-1;\ \Delta,N) - g_2(-1;\ \Delta,N) \geq 0\ . \tag{5.12}$$

Also by the same argument it follows that there exists $\rho^* \equiv \rho^*(\Delta,N)$ such that

$$g_1(\rho;\ \Delta,N) - g_2(\rho;\ \Delta,N) \begin{cases} < 0 & \text{for } \rho < \rho^* \\ = 0 & \text{for } \rho = \rho^* \\ > 0 & \text{for } \rho > \rho^* \end{cases}$$

if

$$g_1(-1;\ \Delta,N) - g_2(-1;\ \Delta,N) < 0\ . \tag{5.13}$$

The relation (5.12) holds if, and only if, one of the following conditions hold:

(a) $c^* \leq \Delta/2$

(b) $\Delta/2 < c^* < N^{\frac{1}{2}} \Delta/2$ and

$g(\Delta/2) \leq g(N^{\frac{1}{2}} \Delta/2)$.

The relation (5.13) holds if, and only if, one of the following conditions hold:

(c) $\Delta/2 < c^* < N^{\frac{1}{2}} \Delta/2$

and $g(\Delta/2) > g(N^{\frac{1}{2}} \Delta/2)$

(d) $N^{\frac{1}{2}} \Delta/2 \leq c^*$.

In getting the above equivalence relation we have used the facts on the nature of g and the existence of c^*, as derived earlier.

The above facts and evidences lead us to conclude the following:

(i) $h(\rho; \Delta,N)$ is a decreasing function of ρ for fixed Δ and N if either (a) or (b) holds.

(ii) $h(\rho; \Delta,N)$ is an increasing function of ρ on $(-1, \rho^*]$ and a decreasing function of $[\rho^*, 1]$, if either (c) and (d) holds.

The above conclusions may be seen from the graphs of $h(\rho; \Delta,N)$, in table 2, as a function of ρ.

The asymptotic expansions for the PCC as given in Section 4 also support the conclusions when N is large. Table 4 gives the values of c_3 and c_4 such that $g(c_3) = g(c_4)$. Table 4 can be used to find a lower bound of ρ^* as follows:

If for some ρ_0 and a combination of Δ_0 and N_0, (5.9) holds, then let

$$c_3 = \Delta_0[4 + 2(1+\rho_0)/N_0]^{-\frac{1}{2}}$$

and find c_4 from the table such that $g(c_3) = g(c_4)$. Then

$$\rho^* \equiv \rho^*(\Delta_0, N_0) \geq 1 - \Delta_0^2 N_0/(2c_4^2) \quad (5.14)$$

Table 1 presents the graphs of $h(\rho; \Delta,N)$ for fixed N and $\rho = -0.9999$, $\rho = 0$ and $\rho = 0.9999$ as a function of Δ. Table 3 gives the values of $h(\rho; \Delta,N)$ for various values of Δ, N and ρ.

5.3. Estimation of probability of correct classification and Inequalities.

When all the parameters are known, the minimax Bayes rule ψ_0^* is the one which accepts H_1 if, and only if

$$(2X_0 - \mu_1 - \mu_2)' \Sigma^{-1}(\mu_1 - \mu_2) \geq 0 . \quad (5.15)$$

The PCC of this rule under H_1 or H_2 equals $\Phi(\Delta/2)$. Several estimates of $\Phi(\Delta/2)$ are proposed in the literature. In effect, these estimates are considered in order to ascertain the optimum PCC under complete knowledge and give some information on the divergence between the two populations. Following Fisher [12] and Smith [25] we consider two such estimates, given by $\Phi(\hat{\Delta}/2)$ and

$c_1(\psi_0)$ respectively, where

$$(5.16) \qquad \hat{\Delta}^2 = ((\bar{X}_1 - \bar{X}_2; B))$$

and $c_1(\psi_0) \equiv$ the proportion of the observations in the training sample correctly classified by ψ_0 as coming from Π_1, and ψ_0 is given by (5.15).

Also we consider the problem of estimating the PCC of the rule ψ_0 and the above two estimates may also be taken as estimates of the PCC of ψ_0.

In particular, when $p = 1$, $P(\psi_0)$ is easily obtained as

$$\begin{aligned}(5.17) \qquad P(\psi_0) &= P_1[(2X_0 - \bar{X}_1 - \bar{X}_2)(\bar{X}_1 - \bar{X}_2) \geq 0] \\ &= P_1[X_0 - (\bar{X}_1 + \bar{X}_2)/2 \leq 0] + P[(\bar{X}_2 - \bar{X}_1) \leq 0] \\ &\quad - 2\, P_1[X_0 - (\bar{X}_1 + \bar{X}_2)/2 \leq 0,\ \bar{X}_2 - \bar{X}_1 \leq 0] \\ &= \Phi(a) + \Phi(b) - 2\Phi(a)\Phi)b)\end{aligned}$$

where

$$(5.18) \qquad a = (\Delta/2)[1 + (1+\rho)/2N]^{-\frac{1}{2}}$$

and

$$(5.19) \qquad b = (\Delta/2)[(1-\rho)/2N]^{-\frac{1}{2}} .$$

When $p > 1$, the exact expression for $P_1(\psi_0)$ becomes quite involved; however, approximate values are obtained, to the order of n^{-2}, in theorem 4.3.1., theorem 4.3.2. and theorem 4.3.3. according as ψ_0 is equivalent to ψ_{00}, ψ_{10} and ψ_{20} respectively.

First we consider Fisher's estimate given by $\Phi(\hat{\Delta}/2)$. As we have seen in Section 3, for a suitable constant k, $k\hat{\Delta}^2$ has a noncentral chi-square, a noncentral F or a mixture of noncentral F distribution according as B in (5.16) is given by (4.7), (4.8) or (4.9) respectively and hence the exact expression of $E\Phi(\hat{\Delta}/2)$ becomes quite involved. From the expansions in Section 4, we obtain an approximation, to the order of n^{-2}, of $E\Phi(\hat{\Delta}/2)$ as follows:

<u>Case (a)</u>: Σ is known. In this case $B = \Sigma$ and

$$(5.20) \qquad E\Phi(\hat{\Delta}/2) = \Phi(\Delta/2) + n^{-1}\varphi(\Delta/2)[(2\Delta)^{-1}\operatorname{tr}(I_p - \Sigma^{-1}\Sigma_{12}) - \{(2\Delta^3)^{-1} + (4\Delta)^{-1}\}(\mu_1-\mu_2)'(\Sigma^{-1} - \Sigma^{-1}\Sigma_{12}\Sigma^{-1})(\mu_1-\mu_2)](n/N) + O(n^{-2})$$

which reduces to the following when $\Sigma_{12} = \rho\Sigma$:

$$(5.21) \qquad E\Phi(\hat{\Delta}/2) = \Phi(\Delta/2) + n^{-1}\varphi(\Delta/2)[(2\Delta)^{-1}(p-1)(1-\rho) - \Delta(1-\rho)/4](\tfrac{n}{N}) + O(n^{-2}) .$$

<u>Case (b)</u>: Σ is unknown and $\Sigma_{12} = \rho\Sigma$ with known ρ. In this case B is given by (4.8) and

$$(5.22) \qquad E\Phi(\hat{\Delta}/2) = \Phi(\Delta/2) + n^{-1}\varphi(\Delta/2)[(2\Delta)^{-1}(p-1)(1-\rho)(n/N) + (p+1)\Delta^2 - \Delta/8 - \Delta(1-\rho)(n/4N) - \Delta^3/16] + O(n^{-2}).$$

<u>Case (c)</u>: Σ and Σ_{12} are unknown. In this case B is given by (4.9) and

$$(5.23) \qquad E\,\Phi(\hat{\Delta}/2) = \Phi(\Delta/2) + n^{-1}\,\varphi(\Delta/2)[(2\Delta)^{-1}\{tr(I_p - \Sigma^{-1}\Sigma_{12})(n/N)$$

$$+ (p+1)\Delta^2 + (\mu_1-\mu_2)'\,\Sigma^{-1}\Sigma_{12}\Sigma^{-1}\Sigma_{12}\Sigma^{-1}(\mu_1-\mu_2)$$

$$+ (\mu_1-\mu_2)'\,\Sigma^{-1}\Sigma_{12}\Sigma^{-1}(\mu_1-\mu_2)\ tr\,\Sigma^{-1}\Sigma_{12}\}$$

$$- \{(16\Delta^3)^{-1} + (32\Delta^{-1})\}$$

$$\cdot\,[(8n/N)(\mu_1-\mu_2)'(\Sigma^{-1} - \Sigma^{-1}\Sigma_{12}\Sigma^{-1})(\mu_1-\mu_2)$$

$$+ \{\Delta^2 - (\mu_1-\mu_2)'\,\Sigma^{-1}\Sigma_{12}\Sigma^{-1}(\mu_1-\mu_2)\}^2$$

$$+ \{\Delta^2 + (\mu_1-\mu_2)'\,\Sigma^{-1}\Sigma_{12}\Sigma^{-1}(\mu_1-\mu_2)\}^2]] + o(n^{-2})$$

which reduces to the following when $\Sigma_{12} = \rho\,\Sigma$:

$$(5.24) \qquad E\,\Phi(\hat{\Delta}/2) = \Phi(\Delta/2) + n^{-1}\,\varphi(\Delta/2)[(2\Delta)^{-1}(p-1)(1-\rho)n/N$$

$$+ (2p+1)(1+\rho^2)\Delta/8 - \Delta(1-\rho)n/4N$$

$$+ \Delta^3(1+\rho^2)/16] + o(n^{-2})\ .$$

When $p = 1$ and Σ (now a scalar) is known,

$$(5.25) \qquad E\,\Phi(\hat{\Delta}/2) = \Phi(c) + \Phi(-b) - 2G(c, -b, -\rho_1)$$

where

$$(5.26) \qquad c = (\Delta/2)[1 + (1-\rho)/2N]^{-\frac{1}{2}}$$

$$(5.27) \qquad \rho_1 = c/b$$

and b is given by (5.19), $G(\cdot,\cdot,\rho)$ denotes the distribution function of the standard bivariate normal with coefficient of correlation ρ.

Now we consider Smith's estimate $c_1(\psi_0)$. Let, for $\alpha = 1,2,\ldots,N$

$$t_\alpha = \begin{cases} 1 & \text{if } (2X_{1\alpha} - \bar{X}_1 - \bar{X}_2)' B^{-1}(\bar{X}_1 - \bar{X}_2) > 0 \\ 0 & \text{elsewhere} . \end{cases}$$

Then

$$c_1(\psi_0) = (1/N) \sum_{\alpha=1}^{N} t_\alpha$$

and

$$E\, c_1(\psi_0) = E(t_\alpha) .$$

Case (a): In this case Σ is known and $B = \Sigma$. In view of (4.11) we can assume (4.10). Define

$$U_1 = X_{11} - \bar{X}_1$$

and

$$U_2 = \bar{X}_1 - \bar{X}_2 .$$

Then the joint distribution of U_1 and U_2 is given by

$$\begin{pmatrix} U_1 \\ U_2 \end{pmatrix} \sim N_{2p} \left[\begin{pmatrix} 0 \\ \delta \end{pmatrix} , \quad \begin{pmatrix} (1 - \frac{1}{N})I_p & 0 \\ 0 & \frac{2}{N}(I_p - D) \end{pmatrix} \right] .$$

Hence, for $U_0 = U_1' U_2 (U_2' U_2)^{-\frac{1}{2}}$, we have

$$\begin{aligned} E\, c_1(\psi_{00}) &= P[(2U_1 + U_2)' U_2 > 0] \\ &= P[U_0 > -(1/2)(U_2' U_2)^{\frac{1}{2}}] . \end{aligned}$$

Since the conditional distribution of U_0 given U_2 is

$N[0, (1 - 1/N)]$, we have

$$(5.28) \qquad E\, c_1(\psi_{00}) = E\, \Phi[(1 - 1/N)^{-\frac{1}{2}} (U_2' U_2)^{\frac{1}{2}} (1/2)]$$

$$= E\, \Phi[(1 - 1/N)^{-\frac{1}{2}} \hat{\Delta}/2] \;.$$

The distribution of $k\, \hat{\Delta}^2$ is a noncentral chi-square, for some constant k, and so the exact expression of (5.28) is complicated for $p > 1$; however, when $p = 1$, we have

$$(5.29) \qquad E\, c_1(\psi_{00}) = \Phi(d) + \Phi(-b) - 2G(d, -b, -\rho_2)$$

where

$$(5.30) \qquad d = (\Delta/2)[1 - (1+\rho)/2N]^{-\frac{1}{2}}$$

$$(5.31) \qquad \rho_2 = d/b$$

and G is as defined before, b is given by (5.19).

We also note that (5.29) is valid for $p = 1$ even when Σ (now a scalar) is unknown and ρ is known or unknown.

<u>Case (b):</u> Σ is unknown and $\Sigma_{12} = \rho\, \Sigma$ with known ρ. In this case

$$E\, c_1(\psi_{10}) = P[(2X_{11} - \bar{X}_1 - \bar{X}_2)' A^{-1}(\bar{X}_1 - \bar{X}_2) > 0]$$

where A is given by (2.10).

Define

$$U = k\, T^{-1}(\bar{X}_1 - X_{11})$$

$$V = T^{-1}(\bar{X}_1 - \bar{X}_2)$$

where

$$k = (1 - 1/N)^{-\frac{1}{2}}$$

and T is a lower triangular matrix (defined uniquely) such that

$$A = T\,T'.$$

Then

$$E\, c_1(\psi_{10}) = P[U'V \le (k/2)V'V]\ .$$

Note that the distribution of U is free from any parameters and $(\bar{X}_1, \bar{X}_2, A)$ is sufficient and complete. Thus, from Basu's theorem, it follows that U and V are independently distributed.

In view of the transformations given by (4.11), we assume $\Sigma = I_p$. As we have seen in Section 3, for $\alpha = 1,2,\ldots,N$ and for

$$X^*_{1\alpha} = X_{1\alpha} - X_{2\alpha}$$

$$X^*_{2\alpha} = X_{1\alpha} + X_{2\alpha}$$

we have $A = S^*_{11}/2(1-\rho) + S^*_{22}/2(1+\rho)$ where S^*_{ii} is given by (3.6), $i = 1,2$. Define

$$Z_1 = k(\bar{X}^*_1 - X^*_{11})[2(1-\rho)]^{-\frac{1}{2}}$$

and

$$Z_2 = k(\bar{X}^*_2 - X^*_{21})[2(1+\rho)]^{-\frac{1}{2}}\ .$$

Then, we can rewrite A as

$$A = Z_1Z_1' + Z_2Z_2' + S_1 + S_2$$

where Z_1, Z_2, S_1 and S_2 are mutuallly independent

$$Z_i \sim N_p[0, I_p] \qquad i = 1,2$$

and

$$S_i \sim W_p[I_p, N-2] \qquad i = 1,2$$

Now, define

$$Y_1 = [(1-\rho)/2]^{\frac{1}{2}} Z_1 + [(1+\rho)/2]^{-\frac{1}{2}} Z_2$$

and

$$Y_2 = [(1-\rho)/2]^{\frac{1}{2}} Z_1 - [(1+\rho)/2]^{-\frac{1}{2}} Z_2 .$$

Then Y_1 and Y_2 are independent $N_p[0, I_p]$ and

$$A = Y_1Y_1' + Y_2Y_2' + S_1 + S_2 \equiv Y_1Y_1' + S ,$$

where Y_1 and S are independently distributed and

$$S \sim W_p[I_p, 2N - 3] .$$

Note that $T^{-1}Y_1 = U$ and the density of U is given by

$$(\text{const.})(1 - u'u)^{N-2-p/2}, \quad 0 < u'u < 1 .$$

Define

$$W = (U'V)(V'V)^{-\frac{1}{2}} .$$

Then W is distributed as any particular component of U independently of V. The density of W is

$$(\text{const.})(1 - \omega^2)^{\frac{2N-3}{2} - 1}, \quad -1 < \omega < 1 .$$

Let the distribution function of W be F. Then

$$(5.32) \quad E\, c_1(\psi_{00}) = E[F(\tfrac{k}{2}(V'V)^{\frac{1}{2}})] \ .$$

No attempt will be made to evaluate (5.32) exactly, but it will be used to establish inequalities which follow.

Inequalities: For establishing inequalities between PCC's and the expected values of their estimates, Γ is taken as given by (3.1). Since the rules or statistics involved in ψ_0^*, ψ_0, $\Phi(\hat{\Delta}/2)$ and $c_1(\psi_0)$ are invariant under the transformations given by (4.11), without loss of generality, we assume

$$(5.33) \quad \Sigma = I_p,\ \mu_1 = 0,\ -\mu_2 = (\Delta, 0, \ldots, 0) \ .$$

(5.33) can be easily seen by diagonalizing Σ and $(\mu_1-\mu_2)(\mu_1-\mu_2)'$ instead of Σ and Σ_{12} (which is $\rho\,\Sigma$ now) and follow the proof of (3.10).

Since $P_1(\psi_0^*) = P_2(\psi_0^*)$, $P_1(\psi_0) = P_2(\psi_0)$ and ψ_0^* is a unique Bayes rule when the parameters are known, it can be shown following an argument of Hills [13] that

$$(5.34) \qquad \Phi(\Delta/2) > P(\psi_0) \ .$$

In the standard case (i.e., when $\rho = 0$) for $p = 1$, Hills [13] has established the inequality

$$(5.35) \quad E\, c_1(\psi_0) > \Phi(\Delta/2)$$

which was generalized for $p > 1$ by Das Gupta [11]. The inequality (5.35) is established when Σ is known or ρ is known, in our case.

The inequality

$$(5.36) \qquad E\,\Phi(\hat{\Delta}/2) < E\,c_1(\psi_0)$$

is also established when at least one of Σ and ρ is known. The inequality (5.36), when $\rho = 0$, was proved by Hills [13] for $p = 1$ and generalized by Das Gupta [11] for $p > 1$.

Case (a): Σ is known.

From (5.28), it follows immediately that

$$(5.37) \qquad E\,\Phi(\hat{\Delta}/2) < E\,c_1(\psi_{00}) \;.$$

Consider $p = 1$. Now

$$E\,c_1(\psi_{00}) = P[U_1U_2 \geq 0]$$

where

$$U_1 = X_{11} - (\overline{X}_1 + \overline{X}_2)/2$$

and

$$U_2 = (\overline{X}_1 - \overline{X}_2)/2 \;.$$

Since the conditional distribution of U_1, given U_2, is normal with the conditional mean U_2, we have

$$P[U_1 \geq 0 | U_2 \geq 0] > P[U_1 \leq 0 | U_2 \geq 0] \;.$$

Equivalently

$$(5.38) \qquad P[U_1 \geq 0,\ U_2 \geq 0] > P[U_1 \leq 0,\ U_2 \geq 0] \;.$$

Hence,

$$(5.39)\qquad E\ c_1(\psi_{00}) = P[U_1U_2 \geq 0]$$
$$= P[U_1 \geq 0,\ U_2 \geq 0] + P[U_1 \leq 0,\ U_2 \leq 0]$$
$$> P[U_1 \leq 0,\ U_2 \geq 0] + P[U_1 \leq 0,\ U_2 \leq 0]$$
$$= P[U_1 \leq 0] = \Phi(d) \geq \Phi(\Delta/2)$$

where d is given by (5.30). To prove (5.35) for $p > 1$, let ℓ be a $p \times 1$ vector such that $\ell'\ell = 1$. Transforming the p-variate problem to a univariate problem, it follows from (5.28), (5.33) and (5.39) that

$$\Phi(|\ell_1\Delta|/2) < E\ \Phi[k|\ell'(\overline{X}_1 - \overline{X}_2)|/2]$$
$$\leq E\ \Phi[k \sup_{\ell'\ell=1} |\ell'(\overline{X}_1 - \overline{X}_2)|/2]$$
$$= E\ \Phi[k\ \hat{\Delta}/2] = E\ c_1(\psi_{00})$$

where $\ell' = (\ell_1, \ell_2, \ldots, \ell_p)$. Now letting $\ell_1 = 1$, (5.35) is proved for $p > 1$. Hence, combining (5.34) and (5.35) we get

$$(5.40)\qquad E\ c_1(\psi_{00}) > \Phi(\Delta/2) > P_1(\psi_{00})\ .$$

For $p = 1$, it follows from (5.25) that

$$\lim_{\Delta\to 0} E\ \Phi(\hat{\Delta}/2) = 1 - 2G(0,0,-\rho_1)$$

where ρ_1 is given by (5.27) and hence

$$(5.41)\qquad \lim_{\Delta \to 0} E\ \Phi(\hat{\Delta}/2) > 1/2\ .$$

For $p > 1$, as $\Delta \to 0$, $\hat{\Delta}^2$ is proportional to a (central) Chi-squar random variable with p degrees of freedom, and hence the inequality (5.41) is also true for $p > 1$. But, since

$$(5.42) \qquad \lim_{\Delta \to 0} \Phi(\Delta/2) = 1/2$$

we have the following inequality from (5.37), (5.40), (5.41) and (5.42) for sufficiently small Δ and for all p

$$(5.43) \quad E\ c_1(\psi_{00}) > E\ \Phi(\hat{\Delta}/2) > \Phi(\Delta/2) > P(\psi_{00}) .$$

<u>Case (b)</u>: In this case Σ is unknown but ρ is known. We now establish that (5.40) is true in this case also. As we have pointed out before, when $p = 1$,

$$E\ c_1(\psi_{i0}) = E\ c_1(\psi_0), \quad i = 0,1,2$$

which is given by (5.29). Hence it follows from (5.35) that

$$(5.44) \qquad E\ c_1(\psi_{10}) > \Phi(\Delta/2) .$$

For $p > 1$, once again reducing the problem to a univariate problem using a $p \times 1$ vector ℓ such that $\ell'\ell = 1$, it follows from (5.32) and (5.44) that

$$\begin{aligned} E\ c_1(\psi_{10}) &\geq E[F\{(k/2)|\ell'(\bar{X}_1-\bar{X}_2)|(\ell' A\ \ell)^{-\frac{1}{2}}\}] \\ &= E\ \Phi[(k/2)|\ell'(\bar{X}_1-\bar{X}_2|] \\ &> \Phi(|\ell_1\Delta|/2) . \end{aligned}$$

Now letting $\ell_1 = 1$ we have

$$\Phi(\Delta/2) < E\ c_1(\psi_{10}) .$$

Hence

$$(5.45) \quad E\ c_1(\psi_{10}) > \Phi(\Delta/2) > P(\psi_{10}) .$$

<u>Case (c):</u> In this case both Σ and ρ are unknown. No inequality has been established except (5.32) for $p > 1$. However, when $p = 1$, we have

$$(5.46) \quad E\ c_1(\psi_{20}) > \Phi(\Delta/2) > P(\psi_{20}) .$$

6. Admissibility of some classification rules.

In the sequel Γ is assumed to have the structure given by (3.1). It is also assumed that at least one of Σ and ρ is known. Hence, in particular, we have two rules given by (2.8) and (2.9), respectively, for known Σ and unknown Σ but known ρ. These two rules are proved to be admissible Bayes rules. In proving admissibility of the rule given by (2.8), we put unit mass to $\rho = \rho^*$ as the prior distribution of ρ in case ρ is unknown; normal priors are considered for the unknown means. The technique of Kiefer and Schwartz [16], in particular, lemma 3 has been used to prove admissibility of the rule given by (2.9).

Theorem 6.1. When Σ is known and ρ is known or unknown, the class of rules which accepts H_1 if, and only if,

$$((\overline{X}_1 - X_0; \Sigma)) - ((\overline{X}_2 - X_0; \Sigma)) < k$$

for some constant k, is admissible Bayes rule.

Proof: Let $\underline{\overline{X}} = (X_{11}, \ldots, X_{1N}, X_{21}, \ldots, X_{2N}, X_0)$ be the $p \times (2N + 1)$ matrix of observations. The likelihood of $\underline{\overline{X}}$ will be denoted as the conditional probability function, $f(\underline{\overline{X}}, \rho, \Sigma \mid \mu_1, \mu_2)$, given μ_1 and μ_2. Without loss of generality we have assumed ρ known. We can write

$$(6.1)\quad f(\underline{\overline{X}}, \rho, \Sigma \mid \mu_1, \mu_2) = [(2\pi)^{p(2N+1)/2}(1-\rho^2)^{N/2}|\Sigma|^{(2N+1)/2}]^{-1}$$

$$\cdot\ \mathrm{etr}[-(1/2)\Sigma^{-1}\{A + (X_0 - \mu)(X_0 - \mu)'\}]$$

$$\cdot \operatorname{etr}[-(N/2)(1-\rho^2)^{-1}\Sigma^{-1}\{(\bar{X}_1-\mu_1)(\bar{X}_1-\mu_1)'$$

$$+(\bar{X}_2-\mu_2)(\bar{X}_2-\mu_2)'-2\rho(\bar{X}_1-\mu_1)(\bar{X}_2-\mu_2)'\}].$$

Where $\mu = \mu_i$ and $f = f_i$ under H_i $i = 1,2$ and where A is given by (2.10).

Define

$$(6.2) \qquad Y_{10} = (X_0 - \bar{X}_1)(1 + 1/N)^{-1/2}$$

$$(6.3) \qquad Y_{11} = (X_0 + N\bar{X}_1)(1 + N)^{1/2}$$

and

$$(6.4) \qquad Y_{12} = (\bar{X}_2 - \rho\,\bar{X}_1)(1-\rho^2)^{-1/2}\,N^{1/2}\,.$$

Then Y_{10}, Y_{11} and Y_{12} are mutually independent p-variate normal with the same covariance matrix Σ and, under H_1, $E\,Y_{10} = 0$

$$(6.5) \qquad E\,Y_{11} = \mu_1(1+N)^{1/2} \equiv \nu_1, \quad \text{say}$$

$$(6.6) \qquad E\,Y_{12} = (\mu_2 - \rho\mu_1)(1-\rho^2)^{-1}\,N^{1/2} \equiv \nu_2, \quad \text{say}\ .$$

Rewriting (6.1) in terms of Y_{10}, Y_{11} and Y_{12} we have

$$(6.7) \quad f_1(\underline{X}, \rho, \Sigma \mid \nu_1, \nu_2) = [(2\pi)^{(2N+1)p/2}(1-\rho^2)^{N/2}|\Sigma|^{(2N+1)/2}]^{-1}$$

$$\cdot \operatorname{etr}[-(1/2)\Sigma^{-1}\{A + Y_{10}Y_{10}' + (Y_{11}-\nu_1)(Y_{11}-\nu_1)'$$

$$+(Y_{12}-\nu_2)(Y_{12}-\nu_2)'\}]\ .$$

Consider a prior distribution $\pi(\nu_1)$ and $\pi(\nu_2)$ on ν_1 and ν_2

respectively such that ν_1 and ν_2 are independent $N_p[0, \Sigma]$.
Now, since

$$(6.8) \qquad E(Y_{1i}) = E[E(Y_{1i}|\nu_i] = 0 \quad i = 1,2$$

$$(6.9) \qquad \mathrm{Cov}\,(Y_{1i}) = \mathrm{Cov}\,[E(Y_{1i}|\nu_i)] + E[\mathrm{Cov}\,(Y_{1i}|\nu_i)]$$
$$= \Sigma + \Sigma = 2\Sigma$$

we have,

$$(6.10) \qquad f_1(\underline{\overline{X}}, \rho, \Sigma) = \int f_1(\underline{\overline{X}}, \rho, \Sigma\,|\,\nu_1, \nu_2)\, d\Pi(\nu_1)\, d\Pi(\nu_2)$$
$$= [(2\pi)^{p(2N+1)/2}(1-\rho^2)^{N/2}|\Sigma|^{(2N+1)/2}]^{-1}$$
$$\cdot\, \mathrm{etr}[-\tfrac{1}{2}\Sigma^{-1}\{A+Y_{10}Y'_{10} + (1/2)(Y_{11}Y'_{11}+Y_{12}Y'_{12})\}].$$

Similarly, under H_2, define Y_{2i}, $i = 0,1,2$ the same way as Y_{1i}, $i = 1,2,3$ by interchanging $\overline{X}_1$ and $\overline{X}_2$ and considering the same independent normal prior on the means of Y_{21} and Y_{22}, we get

$$(6.11) \quad f_2(\underline{\overline{X}}, \rho, \Sigma) = [(2\pi)^{p(2N+1)/2}(1-\rho^2)^{N/2}|\Sigma|^{(2N+1)/2}]^{-1}$$
$$\cdot\, \mathrm{etr}[-\tfrac{1}{2}\Sigma^{-1}\{A + Y_{20}Y'_{20} + (1/2)(Y_{21}Y'_{21} + Y_{22}Y'_{22})\}]\ .$$

Hence a Bayes rule is to accept H_1 if, and only if,

$$(6.12) \quad f_1(\underline{\overline{X}}, \rho, \Sigma)/f_2(\underline{\overline{X}}, \rho, \Sigma) > k^*$$

for some constant k^*. But, since

$$(6.13) \quad A + Y_{10}Y'_{10} + Y_{11}Y'_{11} + Y_{12}Y'_{12} = A + Y_{20}Y'_{20} + Y_{21}Y'_{21} + Y_{22}Y'_{22}$$
$$= A + X_0X'_0 + [\overline{X}_1\,\overline{X}'_1 + \overline{X}_2\,\overline{X}'_2 - 2\rho\,\overline{X}_1\,\overline{X}'_2]\,\frac{N}{1-\rho^2}$$

(6.12) is equivalent to

$$(6.14) \quad \operatorname{etr}\{\Sigma^{-1}(1/2)[Y_{11}Y_{11}' + Y_{12}Y_{12}' - Y_{21}Y_{21}' - Y_{22}Y_{22}']\} > k^* .$$

Now

$$Y_{11}Y_{11}' + Y_{12}Y_{12}' - Y_{21}Y_{21}' - Y_{22}Y_{22}'$$

$$= (\tfrac{N}{N+1})\{(\overline{X}_2 - X_0)(\overline{X}_2 - X_0)' - (\overline{X}_1 - X_0)(\overline{X}_1 - X_0)'\} .$$

Hence, (6.14) is equivalent to

$$(1/2)(\tfrac{N}{N+1})\{((\overline{X}_2 - X_0; \Sigma)) - ((\overline{X}_1 - X_0; \Sigma))\} > \log k^*$$

which is equivalent to the rule given by (2.8), thus proving the theorem.

<u>Theorem 6.2.</u> When Σ is unknown and ρ is known, the class of rules which accepts H_1 if, and only if,

$$\frac{1 + \operatorname{tr} A^{-1} D_{01}}{1 + \operatorname{tr} A^{-1} D_{02}} < k$$

for some constant k, and where D_{0i} is given by (2.3), is admissible Bayes.

<u>Proof:</u> As before,

$$f_1(\underline{X}, \rho \mid \Sigma, \nu_1, \nu_2) = [(2\pi)^{(2N+1)p/2}(1-\rho^2)^{N/2}|\Sigma|^{(2N+1)/2}]^{-1}$$

$$\operatorname{etr}[-\tfrac{1}{2}\Sigma^{-1}\{A + Y_{10}Y_{10}' + (Y_{11} - \nu_1)(Y_{11} - \nu_1)'$$

$$+ (Y_{12} - \nu_2)(Y_{12} - \nu_2)'\}] .$$

Following Kiefer and Schwartz [16, lemma 3], consider the prior distribution on Σ, ν_1 and ν_2 as follows:

Let η be a $p \times 1$ vector and

$$\Sigma^{-1} = I_p + \eta\eta' \quad \text{with probability } 1$$

$$\Sigma^{-1}\nu_i = \gamma_i \eta \qquad i = 1,2; \text{ with probability } 1.$$

The conditional distribution of γ_1 and γ_2 given η is

$$\begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix} \sim N_2 \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^{-2} & 0 \\ 0 & \sigma^{-2} \end{pmatrix} \right]$$

where

$$(6.15) \qquad \sigma^2 = 1 - \eta'(I_p + \eta\eta')^{-1}\eta = |I_p + \eta\eta'|^{-1}$$

and the density function of η is given by

$$\frac{d\pi(\eta)}{d\eta} = c|I_p + \eta\eta'|^{-(2N-1)/2}$$

$$= c(1+\eta'\eta)^{-(2N-1)/2}$$

where $\frac{d\pi(\eta)}{d\eta}$ is integrable over R^p if $p < (2N-1)$ and c is determined from

$$\int d\pi(\eta) = 1 \; .$$

Hence the conditional density of $\underline{\bar{X}}$, γ_1 and γ_2 given η is

$$(6.16) \qquad f_1(\underline{\bar{X}}, \rho, \gamma_1, \gamma_2|\eta) = [(2\pi)^{(2N+1)p/2+1}(1-\rho^2)^{N/2}]^{-1}$$

$$\cdot \; |I_p + \eta\eta'|^{(2N+1)/2} \; (\sigma^2)$$

$$\cdot \operatorname{etr}[-(1/2)(I_p+\eta\eta')\{A + Y_{10}Y_{10}' + (Y_{11} - \nu_1)(Y_{11} - \nu_1)'$$

$$+ (Y_{12} - \nu_2)(Y_{12} - \nu_2)'\}]$$

$$\cdot \exp[-(1/2)\sigma^2(\gamma_1^2 + \gamma_2^2)].$$

From (6.15) we have

$$\sigma|I_p + \eta\eta'|^{1/2} = 1 .$$

Now,

$$\sigma^2\gamma_i^2 = \gamma_i^2 - \gamma_i\, \eta' \nu_i\, , \quad i = 1,2$$

and

$$\operatorname{tr}(I + \eta\eta')(Y_{1i} - \nu_i)(Y_{1i} - \nu_i)'$$

$$= \operatorname{tr}[Y_{1i}\, Y_{1i}' + (\eta' Y_{1i})^2 - 2\gamma_i\, \eta' Y_{1i} + \gamma_i \eta \nu_i'] .$$

Hence (6.16) can be written as

$$(6.17) \quad f_1(\overline{X}, \rho, \gamma_1, \gamma_2|\eta) = [(2\pi)^{(2N+1)p/2+1}(1-\rho^2)^{N/2}]^{-1}$$

$$\cdot |I_p + \eta\eta'|^{(2N-1)/2}$$

$$\cdot \operatorname{etr}[-(1/2)(I_p + \eta\eta')\{A + Y_{10}Y_{10}'\}]$$

$$\cdot \operatorname{etr}[-(1/2)(Y_{11}Y_{11}' + Y_{12}Y_{12}')]$$

$$\cdot \exp[-(1/2)\{(\gamma_1 - \eta' Y_{11})^2 + (\gamma_2 - \eta' Y_{12})^2] .$$

Integrating γ_1 and γ_2 in (6.17) we get

$$(6.18)\qquad f_1(\underline{\bar{X}},\ \rho|\eta) = [(2\pi)^{(2N+1)p/2}(1-\rho^2)^{N/2}]^{-1}$$

$$\cdot\ |I_p + \eta\eta'|^{(2N-1)/2}\ \mathrm{etr}[-\ (1/2)(Y_{11}Y_{11}'+Y_{12}Y_{12}')]$$

$$\cdot\ \mathrm{etr}\ [-\ (1/2)(I_p + \eta\eta')(A + Y_{10}Y_{10}')]\ .$$

Hence,

$$f_1(\underline{\bar{X}},\ \rho,\ \eta) = f_1(\underline{\bar{X}},\ \rho|\eta)\ \frac{d\pi(\eta)}{d\eta}$$

$$= c\ \ [(2\pi)^{(2N+1)p/2}(1-\rho^2)^{N/2}]^{-1}$$

$$\cdot\ \mathrm{etr}\ [-\ (1/2)(Y_{11}Y_{11}' + Y_{12}Y_{12}')]$$

$$\cdot\ \mathrm{etr}\ [-\ (1/2)(I_p + \eta\eta')(A + Y_{10}Y_{10}')]\ .$$

Integrating η we get

$$(6.19)\qquad f_1(\underline{\bar{X}},\ \rho) = c\ \ [(2\pi)^N(1-\rho^2)^{N/2}]^{-1}$$

$$\cdot\ \mathrm{etr}\ [-\ (1/2)(A + Y_{10}Y_{10}'+Y_{11}Y_{11}'+Y_{12}Y_{12}')]$$

$$\cdot\ |Y_{10}Y_{10}' + A|^{-1/2}.$$

Defining Y_{2i}, $i = 0,1,2$ in the same way as before and considering the prior distributions the same way as under H_1, we get, under H_2

$$(6.20)\qquad f_2(\underline{\bar{X}},\ \rho) = c\ [(2\pi)^N(1-\rho^2)^{N/2}]^{-1}$$

$$\cdot\ \mathrm{etr}\ [-\ (1/2)(A+Y_{20}Y_{20}'+Y_{21}Y_{21}'+Y_{22}Y_{22}')]$$

$$\cdot\ |Y_{20}Y_{20}' + A|^{-1/2}.$$

Hence a Bayes rule is to accept H_1 if, and only if,

(6.21) $f_1(\underline{\bar{X}}, \rho)/f_2(\underline{\bar{X}}, \rho) > k^*$

for some constant k^*. (6.21) is equivalent to, by (6.13)

$$\frac{|Y_{10}Y'_{10} + A|^{-1/2}}{|Y_{20}Y'_{20} + A|^{-1/2}} > k^*$$

which is equivalent to

$$\frac{1 + Y'_{10} A^{-1} Y_{10}}{1 + Y'_{20} A^{-1} Y_{20}} < k^{*-2}$$

which is equivalent to (2.9), thus proving the theorem.

APPENDIX

Proof of Lemma 4.1.1.

$$P[|Y_i| > 4(\log n)^{\frac{1}{2}}] = (\tfrac{2}{\pi})^{\frac{1}{2}} \int_{t_n}^{\infty} e^{-\frac{1}{2}x^2} dx$$

where

$$t_n = 4(\log n)^{\frac{1}{2}}[2(n/N)(1-\rho_i)]^{-\frac{1}{2}} .$$

By Mill's ratio inequality

$$P[|Y_i| > 4(\log n)^{\frac{1}{2}}] \le (\tfrac{2}{\pi})^{\frac{1}{2}} e^{-\frac{1}{2}t_n^2} t_n^{-1} .$$

Now

$$e^{-t_n^2/2} t_n^{-1}$$

$$= e^{-8\log n/(n-N)(1-\rho_i)} [4(\log n)^{\frac{1}{2}}\{2(n/N)(1-\rho_i)\}^{-\frac{1}{2}}]^{-1}$$

$$= [n^{8N/n(1-\rho_i)} (\log n)^{\frac{1}{2}}]^{-1}[2(1-\rho_i)(n/N)]^{\frac{1}{2}}(1/4)$$

$$\le [n^2(\log n)^{\frac{1}{2}}]^{-1}(2)^{-\frac{1}{2}} = o(n^{-2}) .$$

Hence

(A.1) $$P[|Y_i| > 4(\log n)^{\frac{1}{2}}] \le o(n^{-2}) .$$

Similarly

(A.2) $$P[|Z_i| > 2(\log n)^{\frac{1}{2}}] \le o(n^{-2}) .$$

Let B_{kij} be the element of B_k in the i^{th} row and j^{th} column, $k = 1,2$.

Now

$$V_{ii} > 2\log n$$

$$\Leftrightarrow (\sqrt{n}/2)[B_{1ii} + B_{2ii} - 2] > 2 \log n$$

$$\Leftrightarrow (n/2)[B_{1ii} + B_{2ii}] > 2\sqrt{n} \log n + n \;.$$

Hence, for $\theta > 0$

$$\begin{aligned} P[V_{ii} > 2 \log n] &\le e^{-\theta(\sqrt{n} \log n + n/2)} \\ &\quad \cdot E\, e^{(\theta/2)(B_{1ii} + B_{2ii})(N-1)} \\ &\le e^{-\theta(\sqrt{n} \log n + n/2)} \\ &\quad \cdot E \left[e^{\theta(N-1)B_{1ii}} + e^{\theta(N-1)B_{2ii}} \right](1/2) \\ &= e^{-\theta(\sqrt{n} \log n + n/2)} (1 - 2\theta)^{-n/4}, \quad 0 < \theta < 1/2. \end{aligned}$$

The last equality follows from (4.6).

Let

(A.3) $\quad \theta = k/\sqrt{n}$.

Fix k and let n be sufficiently large such that $k/\sqrt{n} = \theta < 1/2$. Then

$$P[V_{ii} > 2 \log n] \le e^{-k\sqrt{n}/2} e^{-k \log n} (1-2k/\sqrt{n})^{-n/4} .$$

But, since $n/4 = (\sqrt{n}/2k)k\sqrt{n}/2$, we have

$$P[V_{ii} > 2 \log n] \le \text{Constant } e^{-k \log n}$$

$$= O(n^{-k}) .$$

Similarly $\quad P[-V_{ii} > 2 \log n] \le O(n^{-k})$.

Hence, for $k > 2$,

(A.4) $\quad P[|V_{ii}| > 2 \log n] \le o(n^{-2})$.

Now let $i \neq j$.

$$V_{ij} > 2 \log n$$

$$\Leftrightarrow (n/4)(B_{1ij} + B_{2ij}) > \sqrt{n} \log n .$$

Hence

$$\begin{aligned} P[V_{ij} > 2 \log n] &\le e^{-\theta\sqrt{n} \log n} E\, e^{(n/4)\theta(B_{1ij} + B_{2ij})} \\ &\le e^{-\theta\sqrt{n} \log n} E \left[e^{(n/2)\theta B_{1ij}} + e^{(n/2)\theta B_{2ij}} \right] (1/2) \\ &= e^{-\theta\sqrt{n} \log n} (1 - \theta^2)^{n/4}, \quad 0 < \theta < 1/2 . \end{aligned}$$

The last equality follows from (4.6). Again, for fixed k and for θ as in (A.3) and for sufficiently large n such that $\theta = k/\sqrt{n} < 1/2$, we have

$$P[V_{ij} > 2 \log n] \le e^{-k \log n}(1-k^2/n)^{-n/4} = O(n^{-k}) .$$

Similarly,

$$P[-V_{ij} > 2 \log n] \leq O(n^{-k}) .$$

Hence, for $k > 2$

$$\text{(A.5)} \qquad P[|V_{ij}| > 2 \log n] \leq o(n^{-2}) .$$

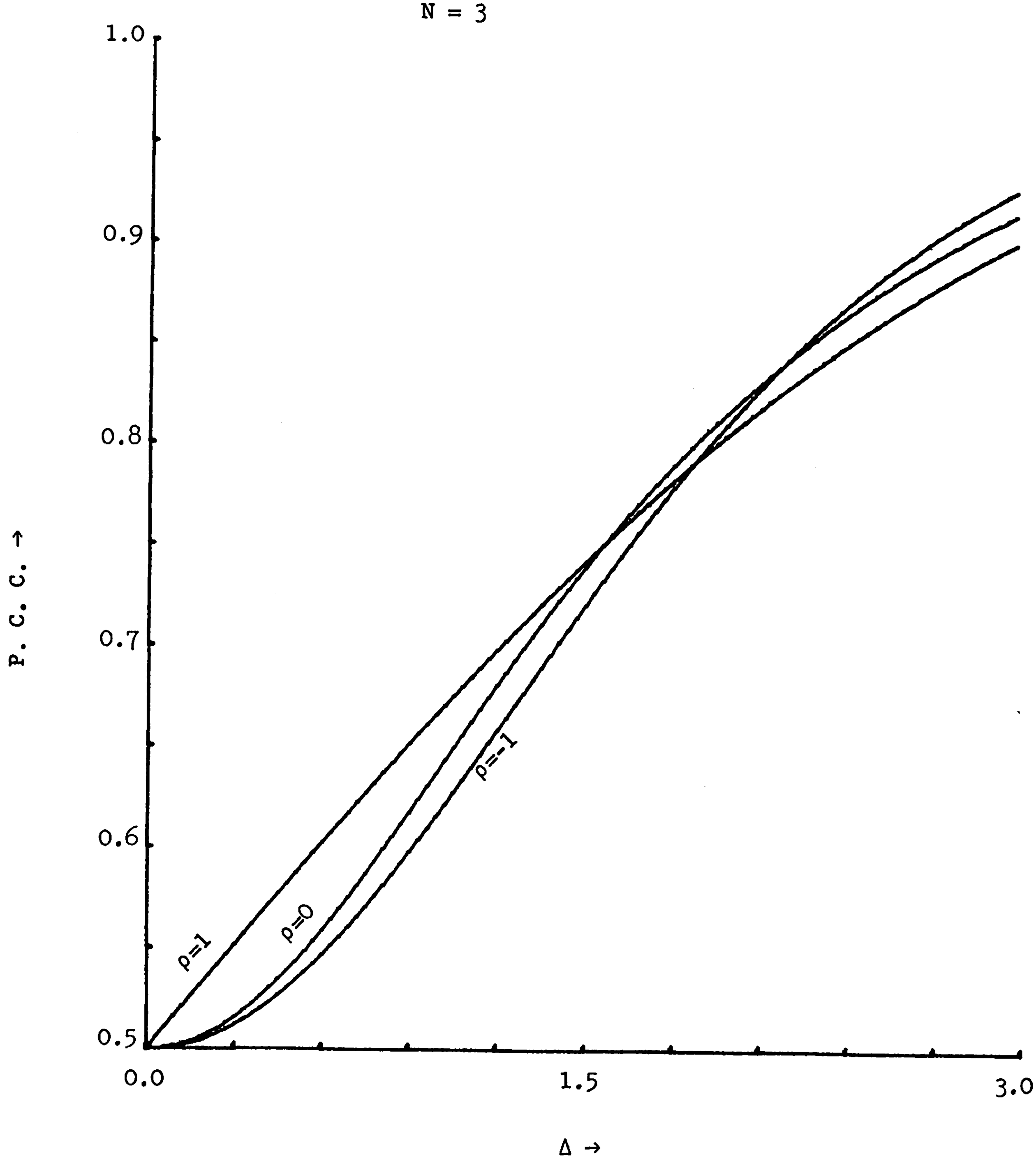
Since, for any finite number of events $E_1, E_2, \ldots, E_m$ we have

$$P[\bigcap_{k=1}^{m} E_k^c] = 1 - P[\bigcup_{k=1}^{m} E_k] \geq 1 - \sum_{k=1}^{m} P(E_k)$$

which when combined with (A.1), (A.2) and (A.5) proves the lemma.

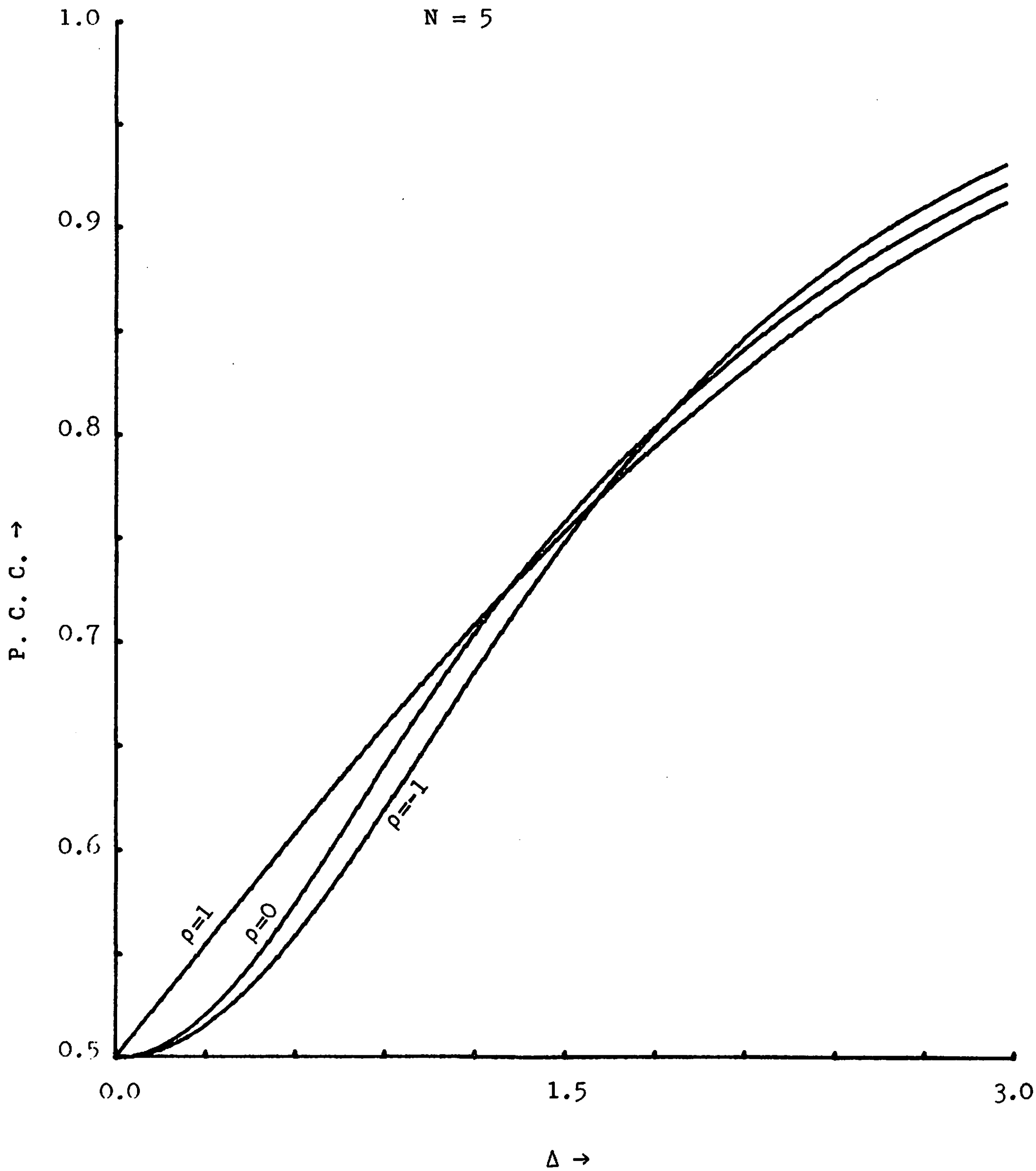
TABLE 1

N = 3



Probability of Correct Classification

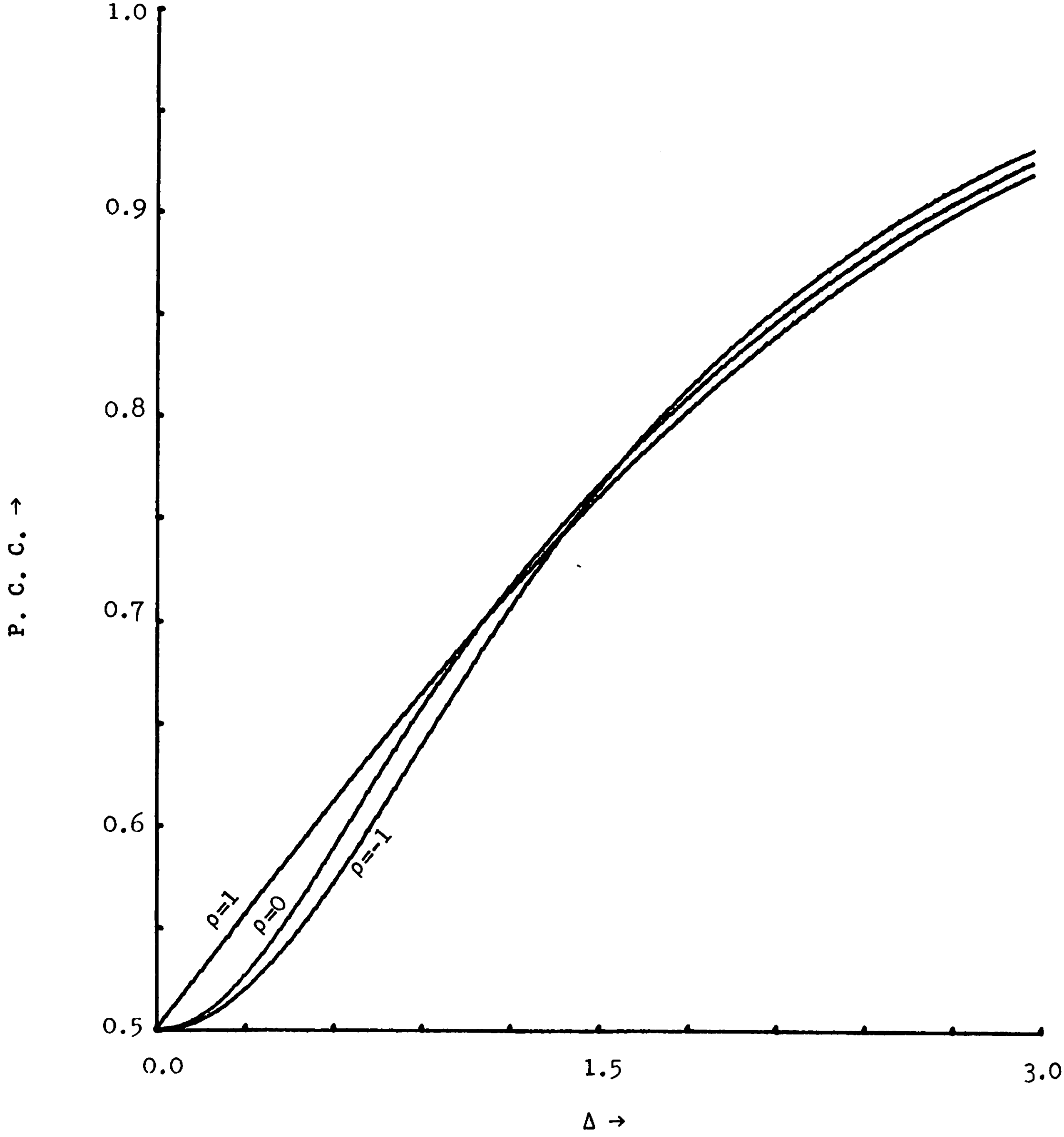
TABLE 1



Probability of Correct Classification

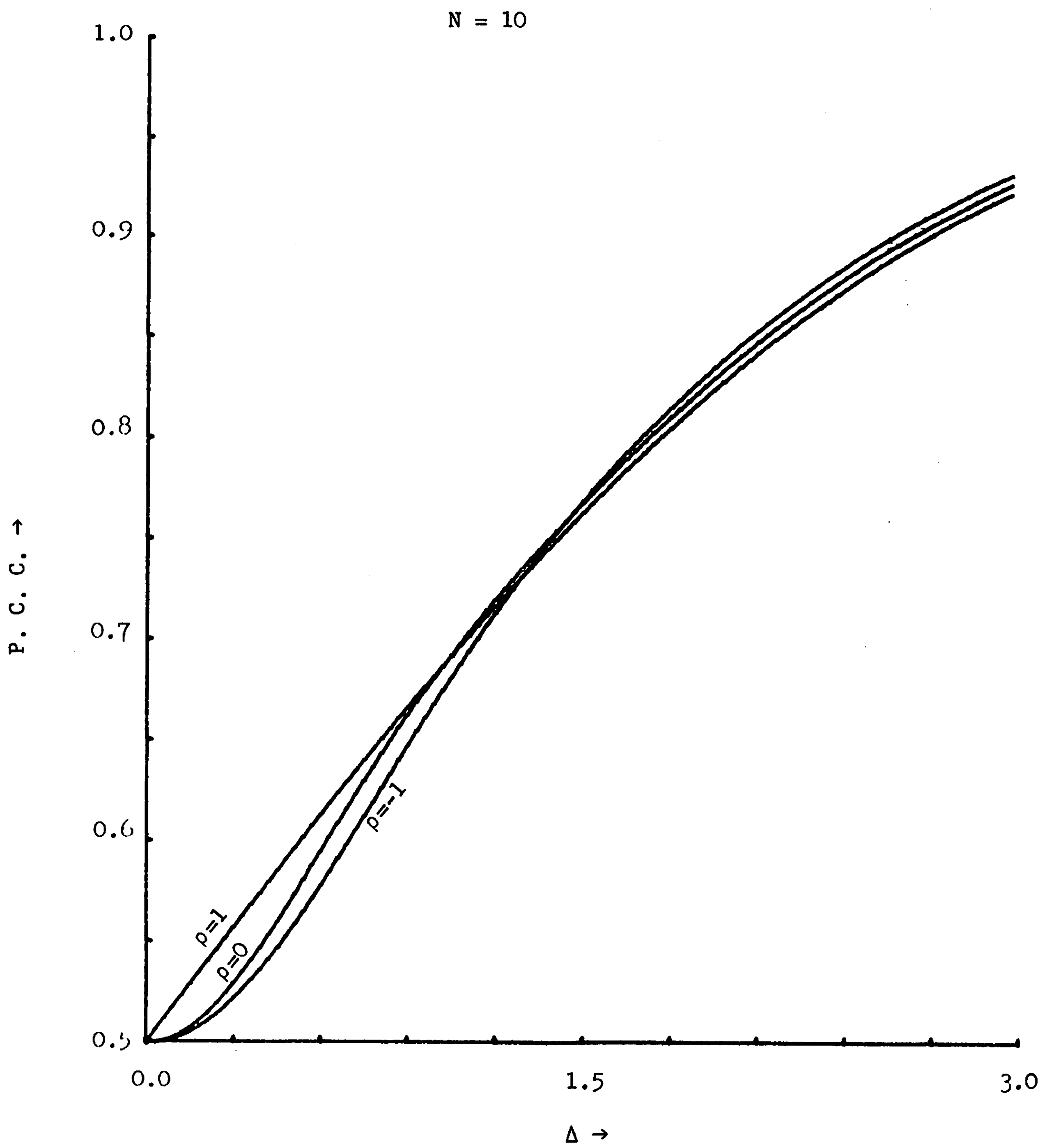
TABLE 1

N = 8



Probability of Correct Classification

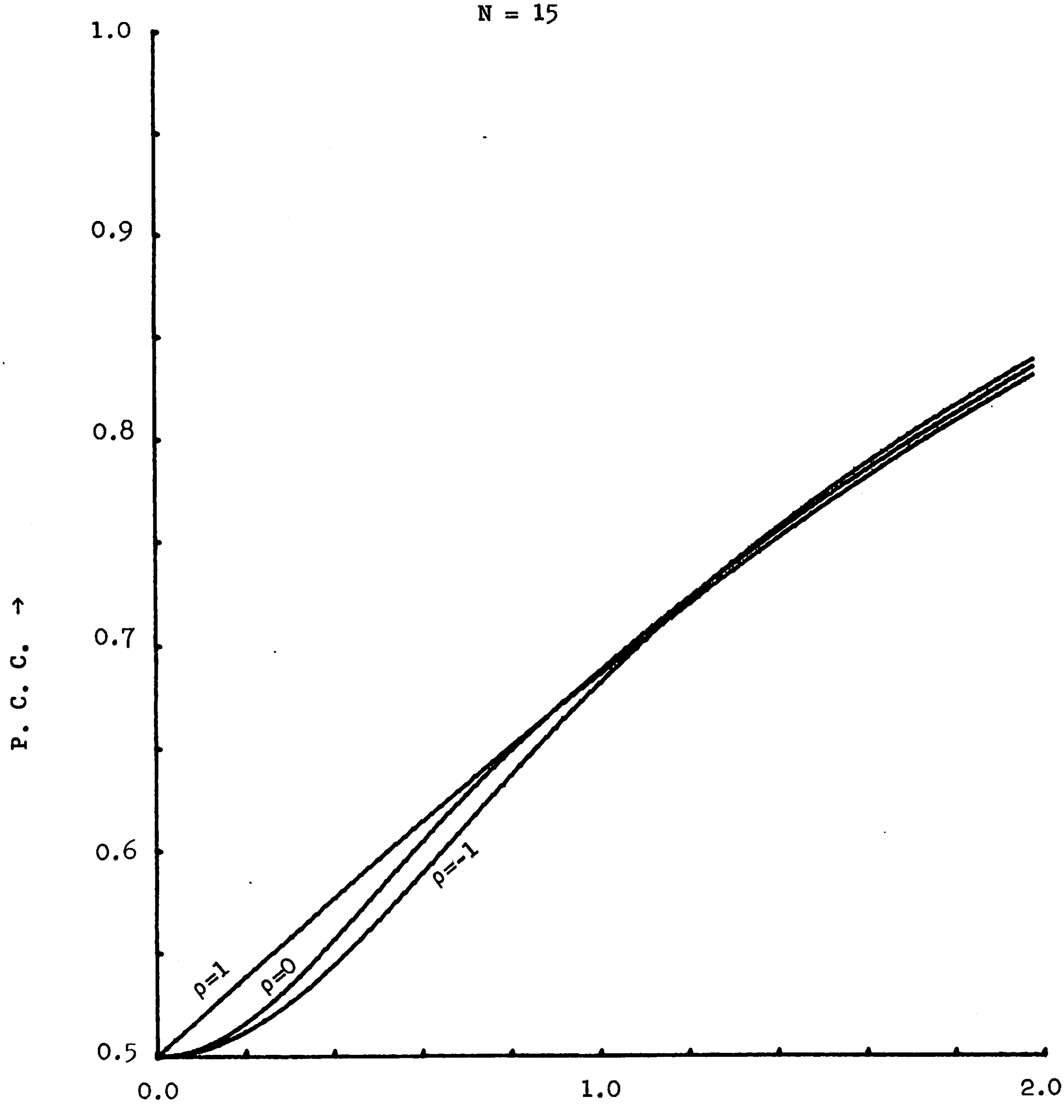
TABLE 1



Probability of Correct Classification

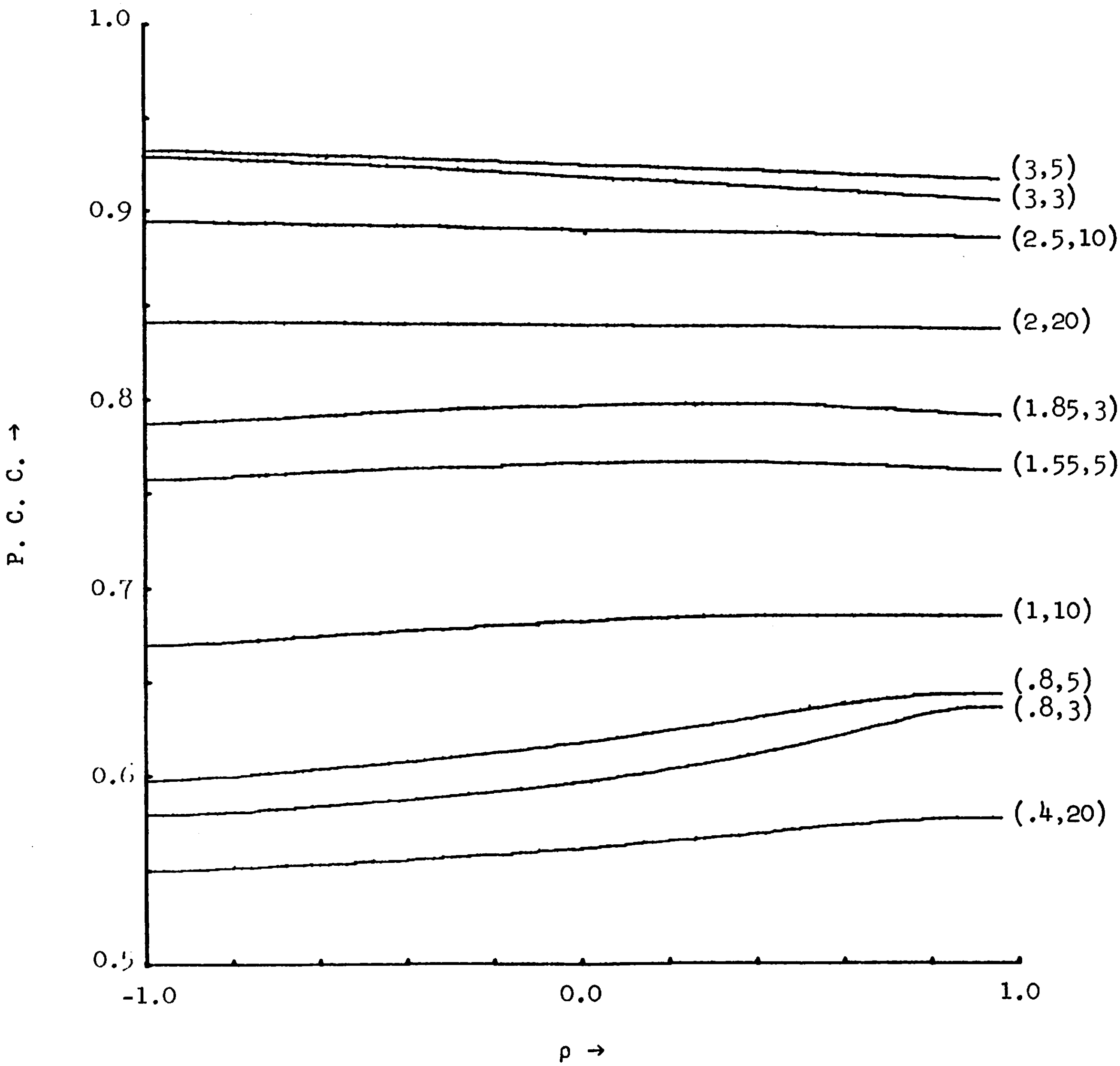
TABLE 1

N = 15



Probability of Correct Classification

TABLE 2



Probability of Correct Classification

for fixed (Δ,N)

TABLE 3

PRØBABILITY ØF CØRRECT CLASSIFICATIØN

DELTA= .25

RHØ N	5	10	15	20	50
-1.00	.51095	.51529	.51849	.52108	.53100
-.90	.51117	.51563	.51890	.52155	.53159
-.80	.51141	.51599	.51934	.52204	.53222
-.70	.51168	.51639	.51982	.52258	.53289
-.60	.51197	.51682	.52034	.52316	.53360
-.50	.51229	.51729	.52091	.52380	.53436
-.40	.51264	.51781	.52153	.52449	.53517
-.30	.51304	.51839	.52222	.52524	.53604
-.20	.51349	.51903	.52298	.52608	.53697
-.10	.51400	.51975	.52383	.52701	.53796
.00	.51458	.52058	.52478	.52804	.53903
.10	.51526	.52152	.52587	.52921	.54019
.20	.51605	.52261	.52712	.53055	.54142
.30	.51701	.52391	.52858	.53208	.54274
.40	.51818	.52547	.53030	.53386	.54413
.50	.51967	.52739	.53233	.53596	.54556
.60	.52162	.52983	.53495	.53847	.54697
.70	.52436	.53308	.53817	.54146	.54821
.80	.52855	.53758	.54224	.54491	.54904
.90	.53593	.54383	.54673	.54800	.54927
1.00	.54542	.54743	.54817	.54855	.54925

TABLE 3

PRØBABILITY ØF CØRRECT CLASSIFICATIØN

DELTA= .50

RHO N	5	10	15	20	50
-1.00	.54184	.55634	.56534	.57269	.59110
-.90	.54260	.55738	.56696	.57381	.59178
-.80	.54343	.55848	.56814	.57497	.59246
-.70	.54433	.55966	.56939	.57619	.59312
-.60	.54532	.56092	.57071	.57747	.59377
-.50	.54639	.56228	.57211	.57881	.59441
-.40	.54756	.56375	.57359	.58021	.59501
-.30	.54883	.56534	.57517	.58167	.59558
-.20	.55033	.56706	.57685	.58320	.59611
-.10	.55195	.56893	.57864	.58478	.59659
.00	.55377	.57098	.58053	.58642	.59701
.10	.55593	.57321	.58254	.58809	.59735
.20	.55819	.57566	.58465	.58979	.59762
.30	.56092	.57334	.58694	.59146	.59781
.40	.56411	.58127	.58908	.59306	.59791
.50	.56739	.59442	.59129	.59449	.59795
.60	.57233	.58772	.59330	.59563	.59793
.70	.57775	.59093	.59451	.59634	.59789
.80	.58399	.59345	.59572	.59656	.59758
.90	.58951	.59437	.59576	.59649	.59726
1.00	.59063	.59420	.59553	.59637	.59775

TABLE 3

PRØBABILITY ØF CØRRECT CLASSIFICATIØN

DELTA= .75

RHØ	N	5	10	15	20	50
-1.00		.58745	.61172	.62477	.63250	.64500
-.90		.58880	.61320	.62606	.63354	.64515
-.80		.59024	.61474	.62738	.63458	.64527
-.70		.59179	.61635	.62872	.63561	.64537
-.60		.59346	.61802	.63008	.63663	.64545
-.50		.59526	.61977	.63146	.63763	.64550
-.40		.59720	.62158	.63284	.63860	.64553
-.30		.59930	.62347	.63421	.63953	.64554
-.20		.60158	.62541	.63557	.64040	.64552
-.10		.60405	.62741	.63689	.64119	.64550
.00		.60674	.62945	.63815	.64190	.64545
.10		.60967	.63149	.63931	.64250	.64540
.20		.61286	.63352	.64035	.64297	.64534
.30		.61631	.63545	.64122	.64329	.64527
.40		.62002	.63732	.64183	.64347	.64520
.50		.62393	.63870	.64228	.64351	.64513
.60		.62787	.63976	.64243	.64343	.64507
.70		.63147	.64027	.64235	.64329	.64500
.80		.63395	.64025	.64216	.64313	.64493
.90		.63446	.63996	.64194	.64297	.64486
1.00		.63395	.63966	.64173	.64280	.64479

TABLE 3

PROBABILITY OF CORRECT CLASSIFICATION

DELTA= 1.00

RHO N	5	10	15	20	50
-1.00	.64100	.66967	.68135	.68661	.69138
-.90	.64268	.67101	.68220	.68708	.69132
-.80	.64446	.67237	.68301	.68750	.69125
-.70	.64633	.67374	.68379	.68789	.69117
-.60	.64830	.67510	.68452	.68822	.69110
-.50	.65038	.67645	.68520	.68850	.69102
-.40	.65257	.67777	.68581	.68872	.69093
-.30	.65487	.67906	.68635	.68889	.69085
-.20	.65727	.68028	.68681	.68899	.69076
-.10	.65977	.68143	.68717	.68903	.69067
.00	.66236	.68246	.68743	.68900	.69059
.10	.66502	.68336	.68758	.68892	.69050
.20	.66769	.68407	.68762	.68880	.69041
.30	.67031	.68459	.68756	.68863	.69033
.40	.67278	.68486	.68741	.68845	.69024
.50	.67493	.68489	.68719	.68824	.69016
.60	.67654	.68471	.68693	.68804	.69007
.70	.67736	.68438	.68666	.68783	.68998
.80	.67727	.68400	.68639	.68762	.68990
.90	.67665	.68361	.68612	.68741	.68981
1.00	.67596	.68322	.68585	.68721	.68973

TABLE 3

PRØBABILITY ØF CØRRECT CLASSIFICATIØN

DELTA= 1.25

RHØ N	5	10	15	20	50
-1.00	.69605	.72276	.73039	.73280	.73401
-.90	.69767	.72356	.73063	.73279	.73391
-.80	.69935	.72432	.73083	.73275	.73381
-.70	.70107	.72504	.73098	.73268	.73371
-.60	.70284	.72570	.73108	.73258	.73361
-.50	.70464	.72629	.73112	.73245	.73350
-.40	.70646	.72680	.73111	.73230	.73340
-.30	.70829	.72723	.73104	.73212	.73330
-.20	.71011	.72755	.73092	.73192	.73320
-.10	.71189	.72776	.73074	.73170	.73310
.00	.71359	.72786	.73053	.73147	.73300
.10	.71517	.72783	.73027	.73124	.73289
.20	.71656	.72769	.72999	.73099	.73279
.30	.71769	.72743	.72968	.73075	.73269
.40	.71848	.72708	.72937	.73051	.73259
.50	.71835	.72666	.72905	.73026	.73249
.60	.71875	.72621	.72873	.73002	.73239
.70	.71824	.72575	.72841	.72977	.73229
.80	.71747	.72599	.72809	.72953	.73219
.90	.71666	.72484	.72773	.72929	.73209
1.00	.71535	.72436	.72746	.72905	.73199

TABLE 3

PRØBABILITY ØF CØRRECT CLASSIFICATIØN

DELTA= 1.50

RHØ N	5	10	15	20	50
-1.00	.74780	.76853	.77237	.77315	.77337
-.90	.74902	.76873	.77221	.77293	.77326
-.80	.75024	.76887	.77202	.77270	.77315
-.70	.75144	.76895	.77181	.77245	.77303
-.60	.75261	.76897	.77156	.77220	.77292
-.50	.75375	.76892	.77129	.77194	.77281
-.40	.75482	.76881	.77100	.77168	.77270
-.30	.75582	.76862	.77069	.77141	.77259
-.20	.75671	.76837	.77036	.77114	.77247
-.10	.75747	.76805	.77002	.77086	.77236
.00	.75807	.76768	.76967	.77059	.77225
.10	.75849	.76725	.76932	.77032	.77214
.20	.75865	.76679	.76896	.77005	.77203
.30	.75857	.76629	.76860	.76977	.77192
.40	.75823	.76579	.76825	.76950	.77180
.50	.75764	.76527	.76789	.76923	.77169
.60	.75685	.76476	.76754	.76896	.77153
.70	.75596	.76424	.76719	.76869	.77147
.80	.75504	.76374	.76684	.76843	.77136
.90	.75412	.76323	.76649	.76816	.77125
1.00	.75322	.76273	.76614	.76789	.77114

TABLE 3

PROBABILITY OF CORRECT CLASSIFICATION

DELTA= 1.75

RHO N	5	10	15	20	50
-1.00	.79363	.80746	.80900	.80918	.80921
-.90	.79426	.80722	.80866	.80890	.80909
-.80	.79484	.80694	.80831	.80861	.80898
-.70	.79536	.80662	.80796	.80832	.80886
-.60	.79581	.80625	.80759	.80803	.80874
-.50	.79619	.80585	.80722	.80773	.80862
-.40	.79645	.80541	.80684	.80744	.80850
-.30	.79661	.80494	.80646	.80715	.80838
-.20	.79663	.80445	.80608	.80686	.80827
-.10	.79652	.80393	.80570	.80657	.80815
.00	.79625	.80340	.80532	.80628	.80803
.10	.79583	.80285	.80494	.80599	.80791
.20	.79524	.80230	.80456	.80570	.80779
.30	.79451	.80175	.80418	.80541	.80768
.40	.79365	.80119	.80380	.80513	.80756
.50	.79271	.80064	.80342	.80484	.80744
.60	.79172	.80010	.80305	.80456	.80733
.70	.79072	.79955	.80267	.80427	.80721
.80	.78974	.79901	.80230	.80399	.80709
.90	.78876	.79847	.80193	.80371	.80697
1.00	.78779	.79794	.80156	.80342	.80686

TABLE 3

PRØBABILITY ØF CØRRECT CLASSIFICATIØN

DELTA= 2.00

RHØ N	5	10	15	20	50
-1.00	.83269	.84081	.84131	.84134	.84134
-.90	.83273	.84034	.84092	.84104	.84122
-.80	.83270	.83985	.84053	.84074	.84110
-.70	.83260	.83934	.84013	.84044	.84098
-.60	.83242	.83881	.83974	.84014	.84086
-.50	.83215	.83827	.83934	.83984	.84074
-.40	.83178	.83772	.83895	.83954	.84062
-.30	.83131	.83715	.83855	.83925	.84050
-.20	.83075	.83659	.83816	.83895	.84033
-.10	.83007	.83601	.83777	.83865	.84026
.00	.82931	.83544	.83738	.83836	.84014
.10	.82845	.83487	.83699	.83806	.84002
.20	.82751	.83430	.83660	.83777	.83990
.30	.82652	.83373	.83621	.83743	.83978
.40	.82550	.83316	.83583	.83713	.83966
.50	.82446	.83260	.83544	.83689	.83954
.60	.82342	.83204	.83506	.83660	.83942
.70	.82239	.83148	.83468	.83631	.83931
.80	.82136	.83092	.83430	.83602	.83919
.90	.82035	.83037	.83392	.83573	.83907
1.00	.81934	.82982	.83354	.83544	.83895

TABLE 3

PRØBABILITY ØF CØRRECT CLASSIFICATIØN

DELTA= 2.25

RHØ	N	5	10	15	20	50
-1.00		.86531	.86957	.86970	.86971	.86971
-.90		.86489	.86901	.86931	.85941	.86959
-.80		.86440	.86845	.86891	.86911	.86947
-.70		.86385	.86789	.86852	.86881	.86935
-.60		.86323	.86732	.86813	.86852	.86923
-.50		.86254	.86674	.86773	.86822	.86911
-.40		.86178	.86617	.86734	.86793	.86899
-.30		.86096	.86559	.86695	.86764	.86887
-.20		.86007	.86502	.86656	.86734	.86876
-.10		.85913	.86444	.86618	.86705	.86864
.00		.85815	.86387	.86579	.86676	.86852
.10		.85713	.86330	.86540	.86647	.86840
.20		.85609	.86274	.86502	.86618	.86828
.30		.85504	.86217	.86464	.86589	.86816
.40		.85398	.86161	.86425	.86560	.86805
.50		.85293	.86105	.86387	.86531	.86793
.60		.85188	.86049	.86349	.86502	.86781
.70		.85084	.85994	.86311	.86473	.86769
.80		.84982	.85938	.86274	.86445	.86758
.90		.84880	.85883	.86236	.86416	.86746
1.00		.84777	.85828	.86198	.86387	.86734

TABLE 3

PROBABILITY OF CORRECT CLASSIFICATION

DELTA= 2.50

RHO N	5	10	15	20	50
-1.00	.89230	.89432	.89435	.89435	.89435
-.90	.89159	.89376	.89397	.89407	.89424
-.80	[illegible]	.89320	.89359	.89379	[illegible]
-.70	[illegible]	.89264	.89321	.89350	.89401
-.60	.88916	.89208	.89284	.89321	.89389
-.50	.88826	.89152	.89246	.89293	.89378
-.40	.88732	.89096	[illegible]	[illegible]	.89367
-.30	.88635	.89040	.89171	.89237	.89355
-.20	[illegible]	.88985	.89134	[illegible]	.89344
-.10	.88434	.88930	.89096	.89130	.89333
.00	.88331	.88874	.89059	.89152	[illegible]
.10	[illegible]	[illegible]	[illegible]	.89124	.89310
.20	[illegible]	.88765	[illegible]	[illegible]	[illegible]
.30	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
.40	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
.50	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
.60	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
.70	[illegible]	[illegible]	.88301	[illegible]	[illegible]
.80	[illegible]	[illegible]	.88765	[illegible]	[illegible]
.90	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
1.00	[illegible]	.88234	[illegible]	[illegible]	[illegible]

TABLE 3

PRØBABILITY ØF CØRRECT CLASSIFICATIØN

DELTA= 2.75

RHØ \ N	5	10	15	20	50
-1.00	.91456	.91543	.91543	.91543	.91543
-.90	.91371	.91490	.91508	.91517	.91533
-.80	.91282	.91437	.91473	.91490	.91522
-.70	.91191	.91384	.91437	.91464	.91511
-.60	.91098	.91331	.91402	.91437	.91501
-.50	.91002	.91279	.91367	.91411	.91490
-.40	.90905	.91226	.91331	.91384	.91480
-.30	.90806	.91174	.91296	.91358	.91469
-.20	.90707	.91122	.91261	.91331	.91458
-.10	.90607	.91070	.91226	.91305	.91448
.00	.90507	.91018	.91191	.91279	.91437
.10	.90407	.90966	.91157	.91253	.91427
.20	.90307	.90915	.91122	.91226	.91416
.30	.90208	.90863	.91087	.91200	.91405
.40	.90109	.90812	.91053	.91174	.91395
.50	.90011	.90761	.91018	.91148	.91384
.60	.89914	.90710	.90983	.91122	.91374
.70	.89817	.90659	.90949	.91096	.91363
.80	.89721	.90608	.90915	.91070	.91353
.90	.89625	.90558	.90880	.91044	.91342
1.00	.89530	.90507	.90845	.91019	.91331

TABLE 3

PROBABILITY OF CORRECT CLASSIFICATION

DELTA= 3.00

RHO N	5	10	15	20	50
-1.00	.93235	.93319	.93319	.93319	.93319
-.90	.93197	.93271	.93287	.93295	.93310
-.80	.93103	.93222	.93255	.93271	.93300
-.70	.93018	.93174	.93222	.93247	.93290
-.60	.92926	.93126	.93190	.93222	.93280
-.50	.92834	.93078	.93158	.93198	.93271
-.40	.92741	.93030	.93126	.93174	.93261
-.30	.92647	.92982	.93094	.93150	.93251
-.20	.92554	.92934	.93062	.93126	.93242
-.10	.92460	.92886	.93030	.93102	.93232
.00	.92367	.92838	.92998	.93078	.93222
.10	.92274	.92791	.92966	.93054	.93213
.20	.92181	.92743	.92934	.93030	.93203
.30	.92089	.92696	.92902	.93006	.93193
.40	.91997	.92649	.92870	.92982	.93184
.50	.91906	.92601	.92838	.92958	.93174
.60	.91815	.92554	.92807	.92934	.93164
.70	.91724	.92507	.92775	.92910	.93155
.80	.91634	.92460	.92743	.92886	.93145
.90	.91544	.92414	.92712	.92858	.93135
1.00	.91455	.92367	.92680	.92835	.9312[illegible]

TABLE 3

PRØBABILITY ØF CØRRECT CLASSIFICATIØN

DELTA= 3.25

RHØ \ N	5	10	15	20	50
-1.00	.94779	.94792	.94792	.94792	.94792
-.90	.94697	.94749	.94763	.94770	.94783
-.80	.94613	.94705	.94734	.94749	.94775
-.70	.94529	.94662	.94705	.94727	.94766
-.60	.94445	.94619	.94677	.94705	.94757
-.50	.94360	.94576	.94648	.94684	.94749
-.40	.94275	.94533	.94619	.94662	.94740
-.30	.94190	.94490	.94590	.94641	.94731
-.20	.94105	.94447	.94562	.94619	.94723
-.10	.94020	.94404	.94533	.94598	.94714
.00	.93935	.94361	.94504	.94576	.94705
.10	.93851	.94318	.94476	.94554	.94697
.20	.93767	.94276	.94447	.94533	.94688
.30	.93683	.94233	.94418	.94511	.94679
.40	.93599	.94190	.94390	.94490	.94671
.50	.93515	.94148	.94361	.94468	.94662
.60	.93432	.94105	.94333	.94447	.94654
.70	.93349	.94063	.94304	.94426	.94645
.80	.93266	.94020	.94276	.94404	.94636
.90	.93184	.93978	.94247	.94333	.94628
1.00	.93102	.93935	.94219	.94361	.94619

TABLE 4

c_3	c_4	$g(c_3) = g(c_4)$
1.6000000	1.9125000	.3119126
1.5000000	1.9722656	.3154487
1.4000000	2.0030823	.3248110
1.3000000	2.0656786	.3408857
1.2000000	2.1302311	.3651493
1.1000000	2.1968008	.3999538
1.0000000	2.3341008	.4490353
.9000000	2.4070415	.5184492
.8000000	2.5574816	.6183791
.7000000	2.7173242	.7668828
.6000000	2.8446987	.9983460
.5000000	3.0224924	1.3848429
.4000000	3.2113982	2.0990144
.3000000	3.4434719	3.6448145
.2000000	3.7293069	8.0647041
.1000000	4.1617827	31.9373044

TABLE 4

c_3	c_4	$g(c_3) = g(c_4)$
.0900000	4.2222571	39.4037921
.0800000	4.2884876	49.8421588
.0700000	4.3617771	65.0674085
.0600000	4.4445721	88.5255923
.0500000	4.5399253	127.4301106
.0400000	4.6529801	199.0498157
.0300000	4.7936288	353.7837730
.0200000	4.9833120	795.8808256
.0100000	5.2878779	3183.2049608
.0090000	5.3322865	3929.8577784
.0080000	5.3813949	4973.6980663
.0070000	5.4364084	6496.2262175
.0060000	5.4990833	8842.0473684
.0050000	5.5721041	12732.5015250
.0040000	5.6599234	19894.4739497
.0030000	5.7707741	35367.8711636
.0020000	5.9228137	79577.5774889
.0010000	6.1725270	318309.9916450

TABLE 4

c_3	c_4	$g(c_3) = g(c_4)$
.00090	6.20945	392975.27344
.00080	6.25043	497359.30226
.00070	6.29652	649612.11741
.00060	6.34923	884194.23260
.00050	6.41093	1273239.64826
.00040	6.48553	1989436.89071
.00030	6.58026	3536776.61211
.00020	6.71117	7957747.24445
.00010	6.92846	31830988.65910
.00009	6.96082	39297516.83662
.00008	6.99679	49735919.72129
.00007	7.03732	64961201.23230
.00006	7.08378	88419412.75343
.00005	7.13829	127323954.31759
.00004	7.20438	198943678.54057
.00003	7.28859	353677650.64126
.00002	7.40546	795774713.86202
.00001	7.60063	3183098853.55350

REFERENCES

[1] Anderson, T. W. (1951). Classification by Mutlivariate Analysis. Psychometrika 16 631-650.

[2] Anderson, T. W. (1958). An Introduction to Multivariate Statistical Analysis. Wiley, New York.

[3] Anderson, T. W. (1971). The Statistical Analysis of Time Series. Wiley, New York.

[4] Anderson, T. W. (1973). Ansymptotic Evaluation of the Probabilities of Misclassification by Linear Discriminant Functions. Discriminant Analysis and Applications (T. Cacoullos, Ed.). Academic Press, New York.

[5] Anderson, T. W. (1973). An Asymptotic Expansion of the Distribution of the Studentized Classification Statistic W. Ann. Statist. 1 964-972.

[6] Bowker, A. H. (1961). A Representation of Hotelling's T^2 and Anderson's Classification Statistic W in Terms of Simple Statistics. Studies in Item Analysis and Prediction (H. Solomon, Ed.). Stanford University Presss.

[7] Bowker, A. H. and Sitgreaves, R. (1961). An Asymptotic Expansion for the Distribution Function of the W-Classification Statistic. Studies in Item Analysis and Prediction (H. Solomon, Ed.). Stanford University Press.

[8] Cacoullos, T. and Styan, P. H. (1973). A Bibliography of Discriminant Analysis, Discriminant Analysis and Applications (T. Cacoullos, Ed.). Academic Press, New York.

[9] Das Gupta, S. (1965). Optimum Classification Rules for Classification into Two Multivariate Normal Populations. Ann. Math. Statist. 36 1174-1184.

[10] Das Gupta, S. (1973). Classification Procedures: A Review. Discriminant Analysis and Applications (T. Cacoullos, Ed.). Academic Press, New York.

[11] Das Gupta, S. (1974). Probability Inequalities and Errors in Classification. To Appear in Ann. Statist. 1.

[12] Fisher, R. A. (1936). Use of Multiple Measurements in Taxonomic Problems. Ann. Engen. 7 179-188.

[13] Hills, M. (1966). Allocation Rules and the Error Rates. J. Roy. Statist. Soc. Ser. B 28 1-31.

[14] John, S. (1960). On Some Classification Statistics. Sankhyā 22 301-316. (Correction: 23 308).

[15] John, S. (1961). Errors in Discrimination. Ann. Math. Statist. 32 1125-1144.

[16] Keifer, J. and Schwartz, R. (1965). Admissible Bayes Character of T^2-, R^2- and Other Fully Invariant Tests for Classical Multivariate Normal Problem . Ann. Math. Statist. 36 747-770.

[17] Kim, D. Y. (1971). Statistical Inference on Constants of Proportionality Between Covariance Matrices. Ph.D. Thesis, Department of Statistics, Stanford University.

[18] Lachenbruch, P. A. (1965). Estimation of Error Rates in Discriminant Analysis. Ph.D. Thesis, University of California, Los Angeles.

[19] Okamoto, M. (1963). An Asymptotic Expansion for the Distribution of Linear Discriminant Function. Ann. Math. Statist. 34 1286-1301. (Correction: 39 1358-1359).

[20] Olkin, I. and Press, S. J. (1969). Testing and Estimation for Circularly Symmetric Model. Ann. Math. Statist. 40 1358-1373.

[21] Rao, C. R. (1952). Advanced Statistical Methods in Biometric Research. Wiley, New York.

[22] Rao, C. R. (1965). Linear Statistical Inference and Its Applications. Wiley, New York.

[23] Sitgreaves, R. (1962). On the Distribution of Two Random Matrices Used in Classification Procedures. Ann. Math. Statist. 23 263-270.

[24] Sitgreaves, R. (1964). Some Results on the Distribution of The W-Classification Statistic. Studies In Item Analysis and Prediction (H. Solomon, Ed.). Stanford University Press.

[25] Smith, C. A. B. (1947). Some Examples of Discrimination. Ann. Engen 13 272-282.

[26] Sorum, M. J. (1972). Estimating the Expected and the Optimal Probabilities of Misclassification. Technometrics 14 935-943.

[27] Wald, A. (1944). On a Statistical Problem Arising In The Classification of an Individual Into One of Two Groups. Ann. Math. Statist. 15 145-162.

[28] Wald, A. (1950). Statistical Decision Functions. Wiley, New York.